

DOCUMENT RESUME

ED 191 726

SE 032 758

AUTHOR Schenberger, Ann Koch
TITLE The Interrelationship of Sex, Visual Spatial Abilities, and Mathematical Problem Solving Ability in Grade Seven. Parts 1, 2, and 3.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Individualized Schooling.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO WRDCIS-TR-387
PUB DATE Jun 76
CONTRACT NE-C-60-3-0065
NOTE 313p.; Report from the Project on Conditions of School Learning and Instructional Strategies. Contains marginal legibility in Appendices.

EDRS PRICE MF01/PC13 Plus Postage.
DESCRIPTORS Cognitive Development; *Cognitive Measurement; Cognitive Processes; Evaluation; *Grade 7; Learning Problems; Learning Theories; *Mathematics Education; *Problem Solving; Secondary Education; Secondary School Mathematics; *Sex Differences; Sex Stereotypes; Student Characteristics; Tests
IDENTIFIERS *Mathematics Education Research; Spatial Ability

ABSTRACT

This three-volume report deals with the hypothesis that males are more successful at solving mathematical and spatial problems than females. The general relationship between visual spatial abilities and mathematical problem-solving ability is also investigated. The research sample consisted of seventh graders. Each pupil took five spatial tests and a test of mathematical problems. The spatial tests were Card Rotations, Form Board 2, and Hidden Figures 2 from the NLSMA battery; Cubes Comparison from the ETS kit; and Space Relations from the Differential Aptitude Test. Problems were chosen to make three subtests of problems differing in amount of spatial or geometric content. Level of mathematics class and IQ were also recorded for each subject. Questions about sex-differences in performance were investigated by using "t" tests; additional refinement came by analysis of variance using IQ and mathematics level data. The relationships between solving spatial problems, drawing diagrams, and solving mathematical problems are explored with a variety of correlational techniques: these include factor analysis, multiple regression analysis, and canonical correlation analysis. (MP)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED191726

TECHNICAL REPORT NO. 387
(Part 1 of 3 Parts)

**the interrelation-
ship of sex, visual
spatial abilities,
and mathematical
problem solving
ability in grade
seven**

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

R. Rossmiller

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

JUNE 1976

WISCONSIN RESEARCH
AND DEVELOPMENT
CENTER FOR
COGNITIVE LEARNING

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY.



SE 032 758

CN000 651
7/3/78

Technical Report No. 387
(Part 1 of 3 Parts)

THE INTERRELATIONSHIP OF SEX, VISUAL SPATIAL ABILITIES,
AND MATHEMATICAL PROBLEM SOLVING ABILITY IN GRADE SEVEN

by

Ann Koch Schonberger

Report from the Project on Conditions of School
Learning and Instructional Strategies

Thomas A. Romberg
Faculty Associate

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

June 1976

JUN 9 1980

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in the University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

WISCONSIN RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING

MISSION

The mission of the Wisconsin Research and Development Center for Cognitive Learning is to help learners develop as rapidly and effectively as possible their potential as human beings and as contributing members of society. The R&D Center is striving to fulfill this goal by

- conducting research to discover more about how children learn
- developing improved instructional strategies, processes and materials for school administrators, teachers, and children, and
- offering assistance to educators and citizens which will help transfer the outcomes of research and development into practice

PROGRAM

The activities of the Wisconsin R&D Center are organized around one unifying theme, Individually Guided Education.

FUNDING

The Wisconsin R&D Center is supported with funds from the National Institute of Education; the Bureau of Education for the Handicapped, U.S. Office of Education; and the University of Wisconsin.

ACKNOWLEDGEMENTS

Academic scholarship is not always the lonely pursuit it is often thought to be. Writing this acknowledgement provided me with an opportunity to recall the many pleasant relationships established during the research and writing of this dissertation.

First, I wish to thank my dissertation committee. Dr. Thomas Romberg, my major professor, first interested me in mathematical problem solving and assisted me greatly in the design, measurement, and data analysis aspects of my study. Dr. Elizabeth Fennema, whose interest in the effects of sex and spatial ability on mathematics achievement inspired my own interest, provided detailed criticism of this manuscript. Dr. Ann Becker contributed an important outside perspective to the study. I am grateful to all of them for their personal, as well as intellectual, support. Dr. Romberg allowed me to complete this study and other degree requirements in a rather unconventional way. Dr. Fennema and Dr. Becker provided examples of academic careers compatible with motherhood. Dr. Thomas Carpenter, Dr. John G. Harvey, and Dr. Mary Montgomery also read the manuscript and offered helpful comments.

Exchanging ideas about mathematical problem solving with other graduate students was useful to me in formulating my own ideas. The work of Ruth Ann Meyer, Diana Wearne, Don Whitaker, and Don Zalewski

was particularly important. In addition, Diana Wearne and Clement Steele took time from their own studies to code some of the problem solving tests.

The staff of the Bangor Public Schools cooperated by providing subjects for the study. Superintendent Wendell Eaton gave permission for the study. Mr. George Voss, Director of Research and Evaluation, made the necessary arrangements and provided some of the data. Mr. Lionel Lord and Mrs. Nancy Luttrell helped with the testing at the Fifth Street School. Thanks also go to the principals, the mathematics teachers, and the seventh-grade students at the Fifth Street and Union Street Junior High Schools who gave their time and cooperation.

I would also like to thank those who assisted me in the data analysis phase of my study. Dr. Jacob Evanson helped plan the analysis and Ms. Linda Junker assisted in the actual programming. The key-punching and computer operating staff of the Wisconsin Research and Development Center for Cognitive Learning provided competent and efficient service. Mrs. Dorothy Egner, also on the Center's staff, did the final typing.

Finally, I acknowledge the contributions of my family: my parents, my husband, and my children. My parents encouraged my intellectual activities and provided me with a first-class education. My children tolerated my absence in good humor. My husband unselfishly gave up his own research time to care for them as well as providing the emotional and financial support without which this project would have been impossible.

TABLE OF CONTENTS

	page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	vii
LIST OF FIGURES.	x
ABSTRACT	xii
Chapter	
I INTRODUCTION	1
The Purposes of the Study	3
Definitions	4
Outline of the Thesis	7
II REVIEW OF RESEARCH ON VISUAL SPATIAL ABILITIES .	9
Introduction	9
Spatial Subfactors.	10
Sex-Related Differences in Visual Spatial Abilities.	22
Another Type of Spatial Test.	28
An Alternate Framework.	34
Causes of Sex-Related Differences in Spatial Abilities.	41
III REVIEW OF RESEARCH ON MATHEMATICAL PROBLEM SOLVING ABILITY.	45
Introduction.	45
Sex-Related Differences in Mathematical Problem Solving Ability.	46
The Relationship Between Mathematical Problem Solving Ability and Visual Spatial Abilities.	65

Chapter		page
IV	DESIGN OF THE STUDY	89
	Introduction	89
	Rationale for This Study	89
	Selection and Description of the Sample.	92
	Selection of the Tests of Visual Spatial Abilities	95
	Construction of the Test of Mathematical Problem Solving	103
	Scoring Procedures	108
V	CARRYING OUT THE STUDY.	110
	Introduction	110
	The Pilot Study and Resulting Changes.	110
	Gathering the Data for the Main Study.	117
	Scoring Procedures	118
VI	DATA ANALYSIS	131
	Introduction	131
	Sex-Related Differences in Performance on the Tests	132
	Relationships Between Visual Spatial Abilities and Mathematical Problem Solving Ability and Sex-Related Differences in These Relationships	150
VII	SUMMARY AND CONCLUSIONS	183
	Introduction	183
	Summary.	183
	Conclusions of the Study	184
	Limitations of the Study	193
	Suggestions for Future Research.	196
	Educational Implications	197
	Conclusion	198
	REFERENCES	200
	APPENDIX A: Sample Items from a Number of Figural Tests . .	215
	APPENDIX B: Tests Used in This Study.	226

	page
APPENDIX C: Descriptive Statistics from Other Administrations of the Tests Used in This Study	248
APPENDIX D: Scatterplots of Two Methods of Scoring the SR-0 Tests.	252
APPENDIX E: Correlation Matrices.	253

Do not film this page

LIST OF TABLES

Table		page
2.1	<u>t</u> Ratios for Sex Differences in Spatial Test Means in the Project Talent Study.	25
3.1	Sex-Related Differences in Applications Scales, National Longitudinal Study of Mathematical Abilities, X-, Y- and Z-Populations.	56
3.2	Sex-Related Differences by Analysis Scales, National Longitudinal Study of Mathematical Abilities, X-, Y- and Z-Populations.	56
3.3	Correlations Among Three Abilities in the CAA Project (1970): Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	67
3.4	Correlations Among Three Abilities in the CAA Project (1971): Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	69
3.5	Correlation Coefficients of Spatial Tests and Analysis Scales in the NLSMA	71
3.6	Per Cents of Total NLSMA Analysis and Application Scales on Which Significant Main Effects of Spatial Variables Were Observed.	74
3.7	Relationships Among Three Abilities in the Project Talent Study: Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	77
3.8	Correlation Coefficients with Concealed Figures and Problem Solving in Students at Swarthmore and Yale	80
5.1	Descriptive Statistics from the Pilot Study Results.	112
5.2	Descriptive Statistics for Several Different Eighth-Grade Groups Taking the Spatial Tests	113
5.3	Textbook and Geometry Instruction for Each of the Mathematics Classes in the Sample.	119
5.4	Agreement over Coders 1, 2 and 3 on Solutions. . . .	125

Table		page
5.5	Agreement over Coders 1, 2, and 3 on Diagrams. . .	127
5.6	Agreement over Coders 1, 2, and 3 on Correct Representations.	129
6.1	Descriptive Statistics and Comparison of Means of Boys, Girls and Both for the Entire Test Battery	134
6.2	Reliability Estimates for Spatial Tests and Problem Scales for Boys, Girls and Both.	136
6.3	Problem Solving Item Analysis: Proportion of Correct Responses for Boys, Girls and Both	138
6.4	Distribution of Boys and Girls in Levels of Mathematics Classes.	140
6.5	<u>F</u> Values Generated by the Sex-by-Level ANOVAs. . .	142
6.6	Distribution of Boys and Girls in IQ Stanines. . .	145
6.7	<u>F</u> Values Generated by the Sex x IQ ANOVAs (Equal Cell Sizes).	147
6.8	Cell Means for the DAT Space Relations Test: Comparison of the Sex-by-IQ Subsample and the Whole Group.	148
6.9	Cell Means for the Cubes Comparison Test: Comparison of the Sex-by-IQ Subsample and the Whole Group.	149
6.10	The Relative Importance of Using Diagrams in Solving Three Types of Problems	154
6.11	Significant Sex-Related Differences in Correlation Coefficients	156
6.12	Correlations Between the Spatial Variables and Problem Solutions Total for the Whole Group. .	158
6.13	The Solutions Factor	161

Table		page
6.14	The Space Factor	163
6.15	The First Drawing Factor (Problem Types A and B) .	165
6.16	The Second Drawing Factor (Problem Type C)	167
6.17	Standardized Regression Coefficients	169
6.18	Standardized Regression Coefficients: Two- and Three-Dimensional Comparison	179
6.19	Canonical Correlation Analysis: Spatial Composite with Problem Solutions Composite	181

LIST OF FIGURES

Figure		page
1.1	Two Representations of a Rectangular Parallelepiped.	6
2.1	Guilford's Structure-of-Intellect Model	14
2.2	Figural Plane of the Structure-of-Intellect Model	14
2.3	A Translation in 2-space.	36
2.4	A Rotation in 2-space	36
2.5	A Reflection in 2-space	36
2.6	A Translation in 3-space	36
2.7	A Rotation in 3-space	36
2.8	A Reflection in 3-space	36
2.9	An Item Typical of the Cubes Test	39
4.1	A Matrix of Spatial Tests	96
4.2	A Set of Card Rotations Items	98
4.3	Two Cubes Comparison Items.	99
4.4	An Item from Form Board 2	100
4.5	An Item from DAT Space Relations.	102
4.6	A Matrix of Spatial Ability Tests Used in This Study	102
4.7	A Hidden Figures 2 Item	103
5.1	Example from the Instructions for the Cubes Comparison Test	115
5.2	An Example of a Figure Scored One for the Existence of a Diagram.	121

Figure		page
5.3	A Difficult Diagram to Score.	123
5.4	Some Solutions for Problem 6.	126
6.1	Sequence of Data Analysis, Part A, Sex-Related Differences in Performance.	133
6.2	Sex-by-Level Interaction: Cell Means on Diagram T ANOVA Using All Data.	143
6.3	Sequence of Data Analysis, Part B	152

THE INTERRELATIONSHIP OF SEX, VISUAL SPATIAL ABILITIES,
AND MATHEMATICAL PROBLEM SOLVING ABILITY IN GRADE SEVEN

Ann Koch Schonberger

Under the supervision of Professor Thomas A. Romberg

The intent of this study was to examine the hypothesis that males are more successful at solving mathematical and spatial problems than females. The relationship between visual spatial abilities and mathematical problem solving ability was also investigated. Specifically five questions were posed.

1. Are there sex-related differences in problem solving ability in mathematics?
2. Are there sex-related differences in any of the measures of visual spatial abilities?
3. What are the relationships between measures of visual spatial abilities and mathematical problem solving ability? Do these relationships differ by sex?
4. Are the relationships between visual spatial abilities and mathematical problem solving ability different for problems requiring differing amounts of spatial knowledge?
5. Is a subject's use of diagrams when solving mathematical problems related to successful solution of mathematical or spatial problems?

The sample consisted of the seventh grade students from one junior high school in Bangor, Maine. Each student took five spatial tests and a test of mathematical problems. The spatial tests were Card Rotations, Form Board 2, and Hidden Figures 2 from the NLSMA battery, Cubes Comparison from the ETS Kit, and Space Relations from the Differential Aptitude Test. The test of mathematical problems was constructed by the author from an item pool developed by Zalewski. Problems were chosen to make three subtests of problems differing in amount of spatial or geometric content. Level of mathematics class and IQ were also recorded for each subject. Questions about sex-related differences in performance were investigated using the t test; additional refinement was made possible by analysis of variance using the data on IQ and level of mathematics class. The relationships between solving spatial problems, drawing diagrams, and solving mathematical problems were explored using a variety of correlational techniques; these included factor analysis, multiple regression analysis and canonical correlation analysis.

Chapter I

INTRODUCTION

The International Women's Year World Conference, held in Mexico City in the summer of 1975, adopted a ten-year plan of action designed to improve the status of the world's two billion women. More specifically, the document dealt with the status of women with respect to education and training, the political process, employment and related economic roles, health and nutrition, the family, population questions, housing, and social services (New York Times, July 3, 1975). Implementation of this plan will make heavy demands on educators throughout the world. In each country and in each field of specialization, educators should address themselves to the problems of improving the position of women. What then must American educators, specifically mathematics educators, do in order to allow for women's equal participation in all societal activities?

Men far outnumber women in occupations requiring a high level of mathematical competence. Only hypotheses for this imbalance exist. Some possible causes are sex bias in career counseling, discrimination in admission to specialized schools, and differences in sex-role socialization. In addition,

differences in mathematical ability are often suggested as reasons (Carnegie Commission, 1973). Some have said that while girls may be better computers, boys excel at mathematical reasoning (Jarvis, 1964; Maccoby, 1966). If this is true, mathematical reasoning could be the "critical filter" (Sells, 1973) in the scientific and technical job market, since in those occupations the application of mathematics to problems is valued more highly than computing skill.

Increased employment of women in scientific and technical occupations is only a part of women's equal participation in modern society. More women are choosing not to marry or to marry and remain financially independent. As they buy and maintain homes, cars, and other products, they need to be able to solve mathematical problems of a practical nature. The traditional female duties of providing food and clothing continue to provide mathematical problems to be solved.

Women should also share equally in the intellectual activities of society, and the importance of problem solving as an intellectual activity is not to be underestimated.

Solving a problem means finding a way out of a difficulty, a way around an obstacle, attaining an aim which was not immediately attainable. Solving problems is the specific achievement of intelligence, and intelligence is the specific gift of mankind [sic]: solving problems can be regarded as the most characteristically human activity [Polya, 1962, p. v.].

Furthermore, solving mathematics problems has been regarded as the essence of human problem solving; psychologists often use mathematical

problems in their general problem solving research (Berry, 1958, 1959; Carey, 1955; Milton, 1957, 1958; Newell & Simon, 1972; Sweeney, 1953).

Mathematics educators stress the value of teaching problem solving.

Our instruction serves to develop the capacity of the human mind for the observation, selection, generalization, abstraction, and construction of models for use in solving problems in other disciplines. Unless the study of mathematics can operate to clarify and to solve human problems, it has indeed only narrow value [Fehr, 1974, p. 27].

Is it true that males are better at solving mathematical problems than females? If so, what other differences between men and women might be involved? In searching for other cognitive abilities which show sex-related differences and which might be related to mathematical problem solving, one might consider visual spatial abilities. The development of sex-related differences in these abilities parallels or precedes their development in mathematics achievement (Fennema, 1974b; Maccoby & Jacklin, 1974). The use of charts, diagrams, and graphs in all branches of mathematics certainly argues for the logic of this connection. The relationship between spatial and mathematical abilities is discussed in detail in Chapter III.

The Purposes of the Study

The purposes of this study are to investigate the following broad questions.

1. Are there sex-related differences in problem solving ability in mathematics?
2. Are there sex-related differences in any of the measures of visual spatial abilities?

3. What are the relationships between measures of visual spatial abilities and mathematical problem solving ability? Do these relationships differ by sex?
4. Are the relationships between visual spatial abilities and mathematical problem solving ability different for problems requiring different amounts of spatial knowledge?
5. Is a subject's use of diagrams when solving mathematical problems related to successful solution of mathematical or spatial items?

Further refinement of the purpose of the study is found in Chapter IV in the section entitled Rationale for this Study.

Definitions

Essential to any study of mathematical problem solving and visual spatial abilities are definitions of those and related terms.

A mathematical problem is a statement which meets three conditions:

- (1) the statement presents information and an objective whose answer is based on that information;
- (2) the objective or answer can be found by translation of the information into mathematical terms and/or application of rules from mathematical areas such as arithmetic, algebra, logic, reasoning, geometry, number theory or topology; and
- (3) the individual attempting to answer the question or attain the objective does not possess a memorized answer or an immediate procedure [Zalewski, 1974, pp. 4-5].

A word of explanation is in order. The third part of the definition serves to set real problems apart from routine applications or exercises. It does, however, introduce the difficulty that what

is a problem for one person may be an exercise for another. As the definition is used in subsequent chapters, an effort has been made to show how this part of the definition was satisfied.

The next definitions are closely related to the first one.

Mathematical problem solving is the process of attaining the objective(s) specified in a mathematical problem. (See Part 1 of previous definition.)

Mathematical problem solving ability is the ability measured by a test of mathematical problems.

Although the author recognizes that to some people the word "ability" implies a genetic or innate component, a functional definition of ability in terms of a test score will be used in this study. Components of different abilities are discussed in the chapters reviewing the related research.

Some background on mode of representation is helpful in understanding the definition of visual spatial abilities and the other definitions related to it. Bruner has said that any problem in any domain of knowledge can be represented in three ways:

- a) by a set of actions appropriate for achieving a certain result (enactive representation),
- b) by a set of summary images or graphics that stand for a concept without defining it fully (ikonic representation),
- c) by a set of symbolic or logical propositions drawn from a symbolic system that is governed by rules or laws for forming and transforming propositions (symbolic representation) [Bruner, 1964, p. 310].

The power of representation, according to Bruner, is to connect matters that seem on the surface to be separate. He thought that this was crucial in mathematics.

While Bruner's categories are of internal representations (in the mind), others studying external representations have built upon Bruner's definitions.

- E: The enactive mode of representation; it is considered to be synonymous with the object mode and is such that the physical characteristics of the exemplar can be felt and manipulated.
- I: The iconic mode of representation; it is synonymous with the picture mode, and is such that the physical characteristics or qualities can be viewed, but cannot be felt or manipulated independent of the medium in which it is presented.
- S: The symbolic mode of representation; it is taken to be a form of words or symbols (usually mathematical) having only ideational relation to the referent. [Heimer & Lottes, 1973; p. 88].

Bruner's definition of the ikonic mode makes a point which is not emphasized in the latter definition and deserves further comment. That is that the images do not define the situation fully. In a picture certain features or attributes of the situation are ignored. Size, for example is usually ignored; color is often absent. As a result of the nature of the media (paper, film, television) in which iconic representations are presented, such representations are at most two-dimensional. An iconic representation of a three-dimensional object either ignores one dimension or presents a drawing "in perspective." The following two representations of a rectangular parallelepiped or box illustrate this difference.

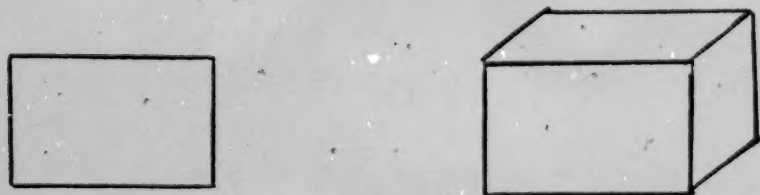


Figure 1.1. Two representations of a rectangular parallelepiped.

The following system which is based on both sets of definitions, is used to categorize external representations of problems in this study.

A concrete representation is one involving objects and is such that the physical characteristics of the exemplar can be felt and manipulated.

A pictorial representation is one in which the physical characteristics can be viewed but cannot be felt or manipulated independent of the medium in which it is presented. Pictorial representations of objects usually disregard some of the objects' attributes.

A symbolic representation is a form of words or symbols having only ideational relation to the referent. There are rules of grammar or logic for combining or transforming such representations.

The following definitions are related to the mode of representation of a mathematical or spatial problem.

A diagram for a problem is a pictorial representation of information presented in the problem or deduced from information in the problem.

Visual spatial abilities are those measured by tests which are recognized in the field of cognitive abilities as spatial and whose stimuli are pictorial representations. According to Werdelin, an aspect common to all such tests is "the ability to comprehend the visual organization of the material and reorganize it" [Werdelin, 1961, p. 77].

A two-dimensional test of visual spatial ability will refer to a test in which the stimuli are planar geometric figures or pictorial representations in which one dimension has been ignored.

A three-dimensional test of visual spatial ability will refer to a test in which the stimuli are pictorial representations in which all three dimensions have been portrayed (drawn in perspective).

Outline of the Thesis

In Chapter II the literature on the nature of visual spatial abilities and sex-related differences in performance on spatial

tests are reviewed. Chapter III contains a discussion of evidence that males and females differ in their ability to solve mathematical problems; the relationship between visual spatial abilities and mathematical problem solving ability are also examined. Chapter IV presents the rationale for this study and the design of the study. In Chapter V the pilot test and changes resulting from it are discussed as well as the gathering of data for the main study and detailed procedures for scoring the tests. Chapter VI deals with the data analysis and its results. In Chapter VII one finds the conclusions and limitations of this study as well as its implications for educational practice and further research.

Chapter II

REVIEW OF RESEARCH ON VISUAL SPATIAL ABILITIES

Introduction

According to Fruchter (1954), spatial abilities have been considered an important part of the general area of aptitudes ever since Galton's work on imagery. As aptitude tests began to be used to predict success in various occupations, tests of spatial ability were developed for the practical purpose of selecting students for schools or training programs in engineering, mechanical arts, or aviation (Fruchter, 1954; Smith, 1964), as well as for theoretical purposes. These tests of spatial abilities are numerous and varied. Some are paper-and-pencil; others require special apparatus. In some tests aural or tactile perception are used, but most tests present visual stimuli. Since the visual tests are the ones which have been linked to mathematical problem solving, further discussion will focus on them. Also the emphasis will be on paper-and-pencil tests. Sample items from many of the tests mentioned in this chapter are found in Appendix A. There are samples of the tests used in this study in Appendix B; all of these tests are of types referred to in this chapter.

The first topic in this chapter is the division of the Space factor into subfactors. Sex-related differences in factor structure and in mean performance on spatial tests (including tests of visual disembedding) are then discussed. The author next presents a theoretical framework to account for these differences. Finally there is a short review of possible causes of the observed sex differences in visual spatial abilities.

Spatial Subfactors

The introduction of factor analytic methods into the study of aptitudes and abilities has supported the existence of a spatial factor but has raised other questions. Is visual spatial ability a unitary trait or are there several such abilities? If there are, in fact, different visual spatial abilities, how should the factors be described and what tests define them?

There was early evidence that there was more than one space factor. During World War II, in the Aviation Psychology Program (Guilford, Fruchter, & Zimmerman, 1952), two factors were definitely isolated, one called visualization and the other called spatial relations. L. L. Thurstone (1950) identified three factors having to do with visual orientation in space. The factor S_1 was described as the ability to identify an object seen from different angles or to visualize a rigid configuration moved into different positions. Thurstone's Flags, in which the subject must decide which flags are only rotations (turned

around) of a referent flag and which ones are reflections (turned over), was given as an example of an S_1 test. Surface Development, in which the subject has to choose which of a set of solids is the one which would result from folding and pasting a paper pattern, was identified as an S_2 test. The factor S_2 was thought to represent the ability to imagine movement within the figure or among parts of the figure. The factor S_3 supposedly involved the body orientation of the observer as an essential part of the problem.

In 1957 Michael, Guilford, Fruchter, and Zimmerman synthesized the previous research and labeled three factors which seemed to them to correspond to those found by the Aviation Psychology Program, by the Thurstones, and also by French in his extensive review of factor analytic studies (1951). This synthesis considered complexity of stimuli, amount of manipulation involved, movement of parts versus movement of the whole, the subject's body orientation, and the relative importance of speed and power. The synthesis was influential. Michael, Guilford, Fruchter, and Zimmerman were active in writing the factor descriptions and selecting the tests for the Kit of Reference Tests for Cognitive Factors developed under the auspices of the Educational Testing Service and referred to subsequently as the ETS Kit. The factors described in the ETS Kit are basically those of the Michael, et al. synthesis, and some subsequent studies (such as the National Longitudinal Study of Mathematical Abilities) followed that framework. For this reason it should be described in some detail.

The first factor, Spatial Relations and Orientation (SR-O), was described as the ability to comprehend the arrangement of elements within a visual stimulus pattern with the subject's body as a frame of reference. In SR-O tests the entire figure is moved into a different position with parts of the figure remaining related to each other in the same way. Tests which have loaded substantially on this factor are a number of the aviation tests, Thurstone's Flags, Cards, Figures, Lozenges (all like Flags), Cubes, and Paper Puzzles, and also Spatial Orientation by Guilford and Zimmerman. The task in Cubes is to decide whether drawings of two marked cubes are of the same or different cubes. In Spatial Orientation the task is to choose a pair of positions for a boat which would result in the pictured pair of views of the shore over the prow of the boat (Guilford, 1967). This factor purportedly combined Thurstone's S_1 and S_3 . The items are usually quite easy and speed is often important in SR-O tests.

The second factor described by Michael, et al. was essentially Thurstone's S_2 and was called Visualization (Vz). On Vz tests the subject is expected to manipulate mentally one or more objects or parts of a configuration according to relatively explicit directions. The subject then must recognize or draw the new configuration. Stimuli are generally more complex in Vz tests and speed is usually less important. The crucial difference between SR-O and Vz tests, according to Michael, et al., is that in SR-O tests the figure as a whole is rigidly transformed, whereas in Vz tests the figure is broken up into parts and the

parts are transformed. Tests believed to be primarily measures of V_z are versions of Thurstone's Mechanical Movements and Surface Development modified for use in the Aviation Psychology Program, Thurstone's Punched Holes and Form Board, and the Guilford-Zimmerman test Spatial Visualization which features pictures of alarm clocks in various positions. Both Form Board and Punched Holes are two-dimensional tests; the rest are all three-dimensional. The task in Form Board is to fit several geometric shapes together to make a larger shape. Kinesthetic Imagery (K), a third factor, was identified by two tests of Thurstone's which appeared to involve right-left discrimination.

Guilford (1967) fit these spatial factors into a broader framework in his three-dimensional, structure-of-intellect model. (See Figure 2.1.) Each cell of the cube represents a factor or ability described by an operation on certain content with a specified product. The letters along the edges of the cube are abbreviations. For example, the cell marked with an asterisk is called Cognition of Figural Systems and abbreviated CFS. Since all the visual spatial tests have figural content according to Guilford's definition, only the plane determined by the figural category is discussed here. (See Figure 2.2.)

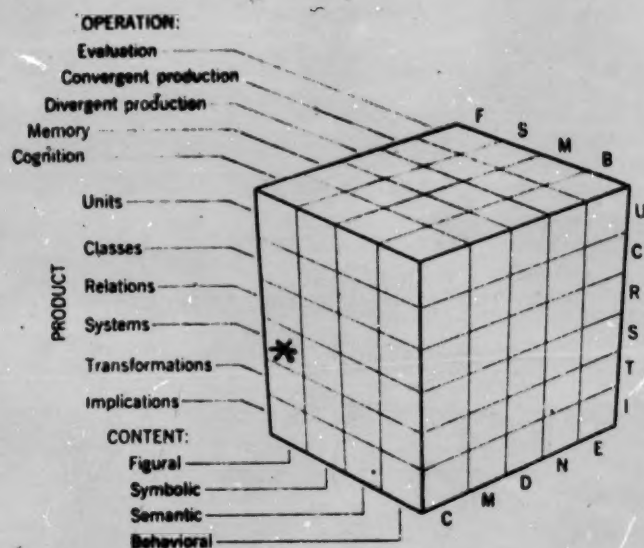


Figure 2.1. Guilford's Structure-of-Intellect Model.

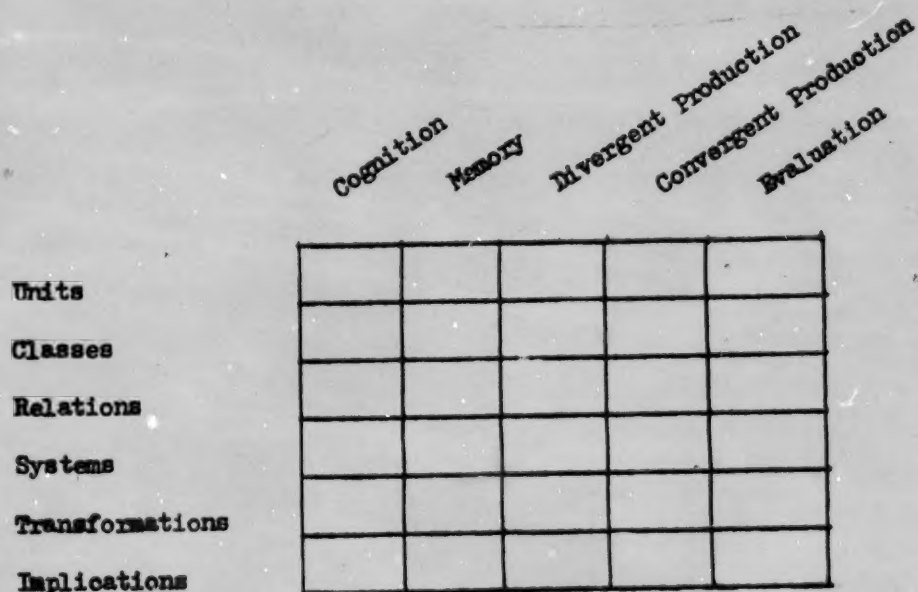


Figure 2.2. Figural plane of the Structure-of-Intellect Model.

According to Guilford, outside of the cognition column most of these abilities have not been demonstrated. Guilford subdivided the CFS cell into three parts--visual, kinesthetic, and auditory. The SR-O factor was identified as CFS-V and the K factor as CFS-K. (Since K tests have not been used to investigate mathematical problem solving ability, they will not be discussed further.) He described CFS-V as the ability to apprehend visually the spatial arrangement of objects with reference to the observer and to one another. The Vz factor was identified as CFT. In other words, SR-O and Vz tests differ only in the product dimension. Transformations of figural material, according to Guilford, involve changes in sensory quantities or qualities, in location, or in arrangement of parts.

The SR-O, Vz framework is not without problems, however, and factor analytic studies have not always shown both factors. Tests were classified according to results of factor analytic studies. A test which has appeared on one factor may satisfy at least part of the description of another factor. For example, Spatial Visualization by Guilford and Zimmerman (1948), classified as Vz, requires the subject to rotate a whole clock in three-dimensional space according to specific instructions. The task seems similar to that in Thurstone's Cubes, an SR-O test. There is often high correlation between SR-O and Vz tests, sometimes almost as high as the reliability of the tests (Smith, 1964). For this reason and because the SR-O, Vz division does not always appear

it may be more accurate to consider SR-O and Vz to be at least two of the subfactors of a single broad space factor.

Both the dimensional and the SR-O, Vz divisions were investigated in Werdelin's (1958) factorial study of mathematical ability. His sample of Swedish high school boys took seven spatial tests adapted from those used in American research. Lozenges and Figures, similar to Thurstone's tests of the same name, were the two-dimensional SR-O tests; Cubes I and Cubes II were the three-dimensional SR-O tests used. Form Boards I and II (the latter similar to the ETS Kit Form Board) test were the two-dimensional Vz tests; the only three-dimensional Vz test used was Solid Figures. The factor analysis yielded only one visual factor, S, on which all seven of the space tests loaded heavily. The fact that the three Vz tests were located at one end of the cluster and the other tests appeared at the other hinted at the two different factors, but Werdelin regarded it as only a faint indication of a subdivision of the space factor.

A subsequent study by Werdelin (1961) investigated the space factor and its relation to geometry in more detail. The subjects were students from three Swedish high schools, two coeducational and one for boys only. The two-dimensional tests used were Figures (SR-O), Form Boards I and II, and a similar test called Puzzles (Vz), as well as a new test called Inversion. In this last test the subject must draw the inverted image of a figure presented on a rectangular grid. Although each figure can be inverted as a whole (SR-O), Werdelin reported that most of his

subjects inverted it piece by piece (Vz). The three-dimensional SR-O tests used were Cubes, which was called Cubes I in the previous study, and Pairs of Cubes, which is like Thurstone's Cubes. Other three-dimensional tests used were Solid Figures (Vz) and a new test called Sections. Items from Sections can be solved either by mentally rotating the whole solid in order to view the section from the perpendicular (which would involve SR-O) or else by analyzing each part of the section (which would involve Vz).

Analysis of the boys' data separately produced a five-factor solution; there were Verbal, Numerical, and General Reasoning factors as well as two space factors identifiable as SR-O and Vz. Werdelin's descriptions of these two factors emphasize the part-whole distinction. When the boys' and girls' data were combined, the two space factors collapsed into one, whereas the other factors remained the same. Separate analysis of the girls' data was not reported.

Werdelin's results indicate that the visual space factor may be different for males and females. Another indication is the fact that many of the studies whose results were synthesized by Michael, et al. had only male subjects. In a study of mathematical ability in college men and women Very (1967) found sex-related differences in the factor structure of spatial ability. Three of the four space tests used had been previously classified as SR-O tests: the ETS Kit tests Cards and Cubes, and the Guilford-Zimmerman test Spatial Orientation. The fourth, Spatial Relations, was attributed to Coleman and was described

as testing the ability to visualize objects in space when rotated. Over half the items on the last test reportedly present no drawings; of the two examples given, one appears to involve rotation and the other asks the subject to determine the change in area of a rectangle whose sides have been doubled. Whether or not this is a spatial test in the usual sense is questionable.

Very's factor analysis with rotation of the complete battery produced an identifiable space factor for both males and females; the major difference between the male and female factors is the high loadings of some reasoning tests on the latter. A second factor with spatial content that appeared for males only includes three reasoning tests and all the spatial tests except Cards. The reasoning test with the highest loading involves identifying changes in the linear order of four letters. The spatial content of this test, Letter Concepts, appears to be as large as the deductive component. In fact Letter Concepts has a significant loading on the women's first space factor. The second factor for males includes spatial tests in which some logic is useful (hence the exclusion of Cards). The first and second male factors seem to be a split of the first female space factor. The third male factor identified by Very as spatial includes only two tests, Space Relations and Practical Estimation, a test requiring judgment of distances, directions, and heights. This factor seems to involve visualizing a spatial configuration given its verbal description. Another factor emerged for women including Practical Estimation, Cards, and two more tests with

figural stimuli. Very neither named nor described this factor. While some of the complexity in these factor structures may be due to factorial impurity of some of the tests, Very's results suggest that there are sex-related differences in the degree to which subjects use reasoning in spatial tests.

The possibility that the visual space factor may be different for males than for females is also suggested by the results of the Concept Attainment Abilities (CAA) Project conducted with a fifth-grade sample (Harris & Harris, 1973). This study used both orthogonal and oblique solutions derived from three different methods of factor analysis to obtain "comparable common factors" of cognitive abilities. A test or variable was defined as "relevant to a comparable common factor," if the variable's loading on that factor exceeded .30 for at least four of the six solutions. Only 2 of the 56 tests used in the 1970 study are like the spatial tests discussed in this chapter: Card Rotations, an SR-O test from the ETS Kit, and Spatial Relations, a form-board type Vz test from the Primary Mental Abilities 4 - 6 battery. Both are two-dimensional. Although one would hardly expect these two tests to define two different factors in such a large battery, it is interesting to note that they did not even appear on the same factor. Spatial Relations is relevant to Comparable Common Factor 6 (CCF 6) for both boys and girls but the loadings are higher for girls. This factor was considered to be the spatial factor. Card Rotations appeared only once out of all 12 of

the solutions for CCF 6. Card Rotations is relevant for boys only on CCF 2, induction of classes when nonsemantic content is used. The girls' loadings for Card Rotations are about equal (between .30 and .40 for two or three of the six solutions) on CCF 2, on CCF 1 which is a verbal factor, and on CCF 7 which involves picture classification.

Unfortunately Card Rotations was not included in the reduced battery of 31 tests given in the 1971 study. However, some of the sex-related differences in the factor structure of the 1971 results appear to be due to Spatial Relations. In the 1971 study CCF 5 was called Simple Visualization and considered to demand visualization of missing portions of figures or pictures. Spatial Relations is relevant to CCF 5 for boys but not for girls. The loadings for this test for girls are about the same for CCF 5 and CCF 3, the Numerical factor. Spatial Relations is relevant to CCF 2 (Nonsemantic Induction) for both sexes.

Not all of the confusion in the visual subfactors can be attributed to sex-related differences in factor structure, however. The SR-O, Vz subdivision has not always appeared even in all-male samples. In a study of problem-solving styles of a sample of men in college or in a college-preparatory high school track French (1965) used four of Thurstone's spatial tests: Cards, Cubes, Punched Holes, and Surface Development. Although French had chosen the tests to allow for two spatial factors, only one appeared.

French's analysis of problem-solving styles may provide some explanation for the ephemeral nature of the spatial subfactors. A factor analytic method for dividing the sample of 177 into two subsamples according to how the subjects solved items yielded 17 pairs of style dichotomies. Data from each of the two subsamples were analyzed separately for each of the 17 divisions. Some of the pairs had similar factor structures; for example, differences in solving items on Cards had no effect on the structure. On the other hand, the division based on Cubes, supposedly also an SR-O test, yielded dramatically different results. The subsample which solved items on Cubes analytically rather than visually had a factor structure in which Cubes no longer appeared with the other spatial tests but with tests which seemed to French to measure induction. This study indicates that subdivision of the visual space factor may be complicated by the fact that some items on spatial tests can be solved in ways which are more logical than visual. Barrett (1953) also reported different styles for solving items from Thurstone's Flags, Figures, and Cards, Space Relations from the Differential Aptitude Test (DAT) which is a surface-development type of test, and Spatial Orientation by Guilford and Zimmerman.

Smith (1964) offered another explanation for difficulty encountered in subdividing the visual space factor. He noted that the tests which mark the Vz subfactor in American research appear together in British factorial studies and have been called the k factor. Smith maintained that the k-loading (Vz loading) of a test "depends on the degree to

which it involves the perception, retention and recognition (or reproduction) of a figure or pattern in its correct proportions" (Smith, 1964, p. 96). According to Smith, this description differentiates Vz or k tests from SR-O tests which require only the ability to manipulate a shape in which proportion remains constant. For example, in solving a form-board item, one has to pay attention to the size and shape of the pieces. For a test to qualify as a k test, success on its items must depend critically on the perception of the figure as a whole. This requirement separates the Vz or k tests from tests of perceptual speed in which the subject must focus on a detail. Whether one accepts Smith's description of the subfactors, or that of Michael, et al., or that of Guilford, the evidence indicates that in most of the factor analyses in which spatial tests split into two groups, the subgroups contain the same tests.

Burt's (1949) review of factor analytic studies further subdivided the space factor into two-dimensional and three-dimensional categories. He cited as additional evidence for this subdivision the frequent comments by teachers (but no empirical research) that success in plane geometry and solid geometry are not highly related. The dimensional division has not been investigated as thoroughly as the SR-O, Vz division, however.

Sex-Related Differences in Visual Spatial Abilities

In addition to the sex-related differences in the factor structure of visual spatial abilities suggested by the work of Werdelin,

Very, and the CAA Project, sex-related differences in mean performance on spatial tests have been noted by a number of authors (Fruchter, 1954; Garai & Scheinfeld, 1968; Maccoby, 1966; Smith, 1964; Tyler, 1965). In order to understand the nature of these sex differences better, one should know at what ages, with which samples, and on which tests they have been observed. One should be cautious in evaluating the significance of results of the large-scale studies. Because of the use of N (the number of subjects) in the denominator of the t ratio (or z ratio), very small differences between means achieve significance. Hayes (1973) offered as a corrective the statistic omega squared (ω^2), interpretable in these studies as the proportion of the variance in a score accounted for by sex.

Maccoby and Jacklin (1974) reviewed more than 20 studies of pre-adolescents published since 1965 using a variety of measures which appeared to have a spatial component. Significant sex-related differences were observed in only three studies, two in favor of girls and one in favor of boys. In the CAA Project's study of fifth-grade students there was no significant sex difference on the V_z test, Spatial Relations, in the 1970 sample. There was a difference in favor of boys significant at the .01 level in the 1971 sample. On Card Rotations, an SR-0 test used only in the 1970 study, the boys' mean was significantly greater than the girls' mean at the .01 level. In a small-scale study of fourth- and sixth-grade students Anglin, Meyer, and Wheeler (1975) used Card Rotations and Form Board as measures of spatial ability. The

latter test was adapted from Werdelin's Puzzles for use in the National Longitudinal Study of Mathematical Abilities and is similar to Space Relations used by the CAA Project. No significant sex differences were observed.

Studies of adolescents have been made using tests of the types discussed earlier in the chapter. Since some have tested subjects in Grades 8 or 9 through 12, these studies will be reviewed one by one instead of by age level. Hobson (1947) reported significant sex differences in the spatial component of the Chicago Tests of Primary Mental Abilities. His subjects were eighth- and ninth-grade students in Brookline, Massachusetts. Although the spatial component appeared to have been measured by several subtests, all by Thurstone, no breakdown of the differences by test was reported.

In the Project Talent survey which attempted to describe high school students in the United States in 1960 (Flanagan, et al., 1964) two tests of visual spatial abilities were given: Two-Dimensional Space, an SR-O test like Thurstone's Figures, and Three-Dimensional Space, a Vz test like Thurstone's Surface Development or the DAT Space Relations. The sample sizes for boys and for girls ranged from about 4000 in the ninth grade to about 3000 in Grade 12. A sex difference was found on each test at each grade level in favor of boys. As might be expected with such large samples, the t ratios are all significant beyond the .01 level. (See Table 2.1.) The differences increase somewhat with age, but this may be a result of the fact that

these were school samples. Since more boys than girls drop out of school in the upper grades, the male population remaining may be more intellectually capable than the female school population. It should be noted with reference to the earlier caution about studies with large N's that for the twelfth grade sample (whose t ratio was one of the largest) sex accounted for 5% of the variance on the Two-Dimensional Space test.

Table 2.1

t-Ratios for Sex Differences in Spatial Test Means
in the Project Talent Study

Test	Grade			
	9	10	11	12
Two-Dimensional Space	12.58	14.45	18.19	17.26
Three-Dimensional Space	11.10	12.19	13.89	14.68

The most recent edition of the manual for the DAT (Bennett, Seashore, & Wesman, 1973) gives norms for boys and for girls in each of the grades from 8 through 12 based on samples of between 5000 and 7000 students. Sex-related differences on the DAT Space Relations Test exceed the .01 level at each grade. The differences increase with age as in the Project Talent study but again there was no control for the fact that more boys, presumably of lower ability, leave school in the upper grades. However, in a study of 204 subjects aged 11 and 14 conducted by Nash (Maccoby & Jacklin, 1974) which used the DAT Space Relations test no sex difference was found among the 11 year olds but a difference in favor of males at age 14 was reported.

In another large-scale study of students in Grades 9 through 12 (Droage, 1967) dropout rate was controlled. The test used was the spatial subtest of the General Aptitude Test Battery which is a type of surface development test (three-dimensional Vz). Sex-related differences of about the same significance as in the Project Talent study were reported at all grade levels but there was little increase in the difference with age. The manual for the Primary Mental Abilities test battery for ages 11 to 17 (T. G. Thurstone, 1958) cites several studies which found sex differences significant at the .01 level on the PMA Space test (Thurstone's Figures). One study was done on 14 year olds in New Zealand; the other was on 16, 17, and 18 year olds in Pittsburgh. To summarize the large-scale studies done primarily in the United States using standardized aptitude tests, it is fair to say that boys' performance is superior to girls' at the .01 level on two-dimensional SR-0 tests and three-dimensional Vz tests. Probably 5% or less of the variance is due to sex. Whether or not these differences really do increase with age is still open to question due to the differential dropout rate in public schools.

In Werdelin's (1961) study of high-ability, Swedish high school students, which involved a number of different spatial tests, sex-related differences appeared on some but not all of the measures. Boys were superior to girls at the .01 level on Solid Figures, Pairs of Cubes, and Sections. Boys from both single-sex and coeducational schools scored significantly better ($p < .05$) than girls on Inversions, and

one group of boys scored higher than girls on Puzzles at the .05 level. There were no significant differences found on Form Board I or II, Cubes, or Figures. Werdelin then took a subsample of 143 male-female pairs who were matched on age, numerical ability, and general reasoning ability. Spatial tests on which boys in the subsample excelled were Sections ($p < .001$), Solid Figures ($.01 < p < .02$), and possibly Pairs of Cubes ($.05 < p < .10$). In trying to differentiate between tests showing sex differences and those which did not, Werdelin generated another way to categorize spatial tests based on the mental processes used in solving the items. He thought that these processes included comprehension of the visual organization of the material, reorganization of the given structures, and manipulation of the space figures. In different tests different aspects are stressed, and the tests on which the boys performed better appeared to Werdelin to stress the reorganization aspect.

Several studies of visual spatial abilities have been done on adolescent subjects and their parents in attempts to investigate the genetic influence on those abilities. Bock and Kolakowski (1973) used the Guilford-Zimmerman test Spatial Visualization form B which involves rotation of a clock in three-dimensional space. As part of their analysis of parent-child correlations they tested all the eleventh-grade students in the two high schools attended by the children in the family study. Sex differences in the eleventh-grade sample of 727 students were significant; sex differences in the family study were not reported.

An earlier study by Stafford (1961) used the Identical Blocks test to investigate parent-child correlations. In this test the subject is required to choose the one of five pictures of three-dimensional solids which is a rotation of the referent solid. This test was not categorized by Michael, et al. or by Guilford. Bock and Kolakowski reported correlations between .5 and .6 for their test and Stafford's with each other and with DAT Space Relations, implying that they are Vz tests, but they cited no references. All three were called tests of spatial visualization. Stafford reported significant sex-related differences in favor of males for both parents and children.

In Very's (1967) study of college students discussed earlier in the chapter, sex-related differences in the means of the spatial tests as well as in the space factor structure were observed. Men's scores on four tests of visual spatial ability were significantly higher than women's at the .01 level. The t values for Spatial Relations (4.83) and Cards (5.82) were lower than those for Cubes (7.31) or Spatial Orientation (7.77). In a study of 50 college men and women Sherman (1974) reported better performance by men ($p < .05$) on the DAT Space Relations test.

Another Type of Spatial Test

Throughout this review certain abilities recognized as spatial have been disregarded because they have shown little relationship to the ability to solve mathematical problems. Among these are abilities

involving aural or tactile perception, mechanical knowledge, or motor skills. On the other hand, there is a group of tests whose classification as spatial has not always been recognized, but whose relationship to mathematical problem solving has often been observed. (This relationship is discussed in the next chapter.) These tests all require the subject to pick a simple figure out of a more complex stimulus. There are a variety of tests of this type and the related ability has a number of names.

Gottschaldt developed the first tests of this factor. In the versions called Gottschaldt A and B which were used by L. L. Thurstone (1944) in his study of perception the subject has to mark or trace the simple figure on top of the complex configuration. In a later adaptation of the Gottschaldt test by Thurstone called Concealed Figures the task is to indicate those complex figures in each row which contain the simple figure at the left. In both Thurstone tests the score is the number right. Guilford's Hidden Figures, also based on the Gottschaldt figures, requires the subject to choose which of five simple figures is contained in each of the complex figures. Witkin's Embedded Figures Test which superimposes color patterns over the simpler achromatic Gottschaldt figures is individually administered and timed. Low score indicates good performance. Two children's forms and a group form of Embedded Figures are available. In the ETS Kit there are two tests of this type. A simple, speeded test, Hidden Patterns, involves indicating each geometric pattern in which a given configuration is embedded. In

Hidden Figures, a more difficult test, the subject must decide which of five geometric figures is embedded in a complex pattern. Hidden Figures was adapted for use in Grades 8 through 12 in the National Longitudinal Study of Mathematical Abilities and called Hidden Figures 2; a simpler version, Hidden Figures 1 was designed for use in Grade 5.

Although tests for this factor are fairly similar, its interpretations and labels differ widely. L. L. Thurstone (1944) first identified a factor involving these tests which he called Gestalt Flexibility in his study of perceptual abilities. Witkin and his followers associated Embedded Figures with two other tests, the Rod and Frame Test and the Body Adjustment Test, and interpreted this factor as a cognitive style labeled Field Independence-Dependence. In the ETS Kit Manual the name of the factor is Flexibility of Closure and it is described as "the ability to keep one or more definite configurations in mind so as to make identification in spite of perceptual distractions" (French, Ekstrom, & Price, 1969, p. 9). In Guilford's structure-of-intellect model these tests were placed in the NFT cell, Convergent Production of Figural Transformations. According to Guilford, the production is convergent because only one answer is acceptable, and the transformation occurs in the form of revision of the use or function of lines.

The classification of these tests as spatial is supported by a number of studies reviewed by Sherman (1967) and in Sherman's own research (1974). Maccoby and Jacklin (1974) followed Sherman in

considering them part of the spatial domain. In their 1974 book they classify the tests as visual spatial (analytic), a change from Maccoby's (1966) categorization of these tests as measures of field independence. The ETS Kit manual (French, Ekstrom, & Price, 1969) notes that its Hidden Figures test has some variance on its version of the SR-0 factor. French found that Concealed Figures had the third highest loading on the spatial factor in his 1965 study, ahead of Punched Holes and Spatial Orientation. Guilford (1967) said that he at first did not differentiate between the CFT (Vz factor) and NFT cells in the structure-of-intellect model but then decided that they differed in operation. To Guilford the Vz tests only require the subject to follow prescribed or resulting operations whereas tests such as Hidden Figures require tearing a new organization out of an old one. Nevertheless, he considered the two types to be close relatives.

The problems in classifying and naming tests of the Gottschaldt-figures type may be due to the fact that they are not factorially pure. Clearly they have a spatial component. The nature and size of the spatial component as well as the identity of the other components is unresolved at present. Also French observed that the subjects in his 1965 study of problem-solving styles used more than one method of solving items on Concealed Figures. For these reasons this author prefers to call them tests of visual disembedding, a term which is descriptive while making the fewest assumptions about the underlying cognitive processes.

Statistically significant sex-related differences in tests of visual disembedding have appeared in a number of studies summarized by Witkin, Dyke, Faterston, Goodenough, and Karp (1962) and have been the subject of a great deal of controversy, although the differences are generally of small magnitude (Kagan & Kogan, 1970). Of the 20 studies using forms of Embedded Figures or Hidden Figures reviewed by Maccoby and Jacklin (1974) there was 1 significant sex difference in favor of females and 8 in favor of males; the remaining 11 showed no significant difference. In the study by Nash summarized by Maccoby and Jacklin sex-related differences on a group form of Embedded Figures paralleled those on DAT Space Relations; differences were not significant at age 11 but were significant at the .01 and .02 levels respectively at age 14. Nash's results might indicate that sex differences in visual disembedding appear during adolescence. In some studies of college students none were found (Jackson, Messick, & Meyers, 1964; Sherman, 1974; Willoughby, 1967), but P. C. Berry (1958) found a significant sex difference in a college sample using Concealed Figures which is a paper-and-pencil test. According to Sherman's review (1967), sex differences in field independence or visual disembedding parallel those in other spatial tests, and she argued that it is the spatial component of the former which accounts for the observed differences.

After such a long and involved catalogue of observed sex-related differences in several types of visual spatial abilities, it is advisable to step back and put these differences into perspective. Even the

biggest differences in means are usually less than half a standard deviation so the within-sex variation is definitely greater than the between-sex variation in spatial skills. Sex-related differences in other abilities are of equal or greater statistical significance: boys' superiority on the DAT Mechanical Reasoning test is higher than it is on Space Relations; girls' superiority on Language Usage or Clerical Speed and Accuracy is as great as boys' superiority on Space Relations (Bennett, Seashore, & Wesman, 1973). Even at the twelfth grade where the sex difference in DAT Space Relations is the greatest, a girl scoring 31 (the girls' mean is 30.9) does better than 45% of the boys.

Also sex-related differences in visual spatial abilities may occur only in certain races, cultures, or levels of general ability. The subjects in the studies surveyed were predominantly white, usually American, and all educated in Western-style schools. Such limitations are not inappropriate to this review because those adjectives describe the subjects to be used in this study. However, it should be noted that similar sex-related differences do not appear in all races or cultures. In particular, none have been observed among Eskimos whose high performance in figural areas has been noted by J. W. Berry (1966) and Kleinfeld (1973). Likewise, none has been found in Canadian Indians or Australian Aborigines (Mitchelmore, 1975). On the other hand, sex-related differences in visual abilities including disembedding have been found in a number of studies of African and Caribbean

populations (J. W. Berry, 1966; Mitchelmore, 1975). However, lack of test-sophistication may be a factor in any one of those results. With regard to ability level one should note that in Werdelin's high-ability Swedish sample sex differences were more limited than those in most studies reviewed here.

An Alternate Framework

Although sex-related differences in visual spatial abilities vary with age and culture and they are no greater than the sex-related differences in some other abilities, some differences do exist. In all but a few groups, men and women perform differently on spatial tasks. However, none of the category systems or descriptions of the tasks presented so far is completely satisfactory in accounting for the sex differences both in mean performance and in factor structure.

Guilford's framework is useful in organizing the large number of tests with figural stimuli. Two of his three categories of spatial ability tests, CFT (Vz) and NFT (visual disembedding), have transformations as their product or output. According to Guilford there had been no precedent in the psychological literature for considering transformations to be products of information. Transformations were considered by Guilford, in the psychological sense, to be changes of any type. Since almost all of the spatial tests involve geometric stimuli, one might instead examine the mathematical transformations involved. In fact, what seems to differentiate tests of visual spatial

abilities from other tests with pictorial stimuli is that they all involve geometric or topological transformations, that is movement of sets of points (figures) in space.

For the purpose of this discussion topological transformations will be divided into two categories--those which are continuous and distance-preserving and those which are not. In nonmathematical language a transformation is continuous if any two points which were close together originally are still close together after the transformation; transformations which include cutting or punching holes are not continuous. A transformation is distance-preserving if the original distance between any two points of a figure is the same as their distance after the transformation; transformations involving folding or moving parts of a figure into different relationships with each other are not distance preserving.

Continuous, distance-preserving transformations, often called rigid transformations, are those which preserve size and shape. There are several different types. In a plane (two-dimensional space) there are three basic types of rigid transformations. In Figure 2.3 the arrow indicates a translation of the triangle a certain distance in a certain direction. In Figure 2.4 the arrow shows the amount and direction of the rotation of the triangle around point P. In Figure 2.5 the triangle is reflected across line L.

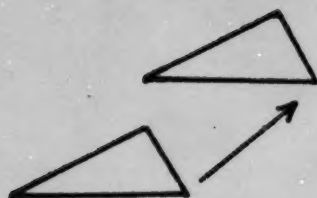


Figure 2.3. A translation in 2-space.

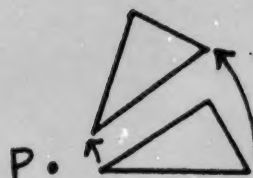


Figure 2.4. A rotation in 2-space.

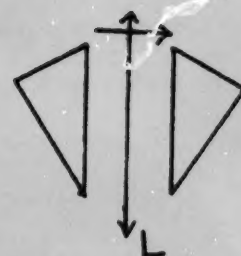


Figure 2.5. A reflection in 2-space.

In three-dimensional space a figure can be translated in a certain direction, rotated about a line, or reflected across a plane. (See Figures 2.6-2.8.)

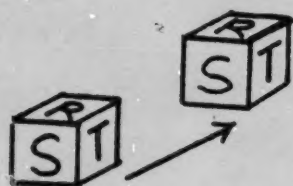


Figure 2.6. A translation in 3-space.

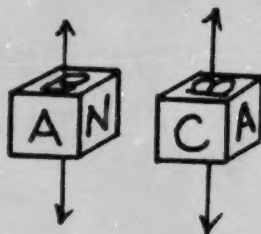


Figure 2.7. A rotation in 3-space.

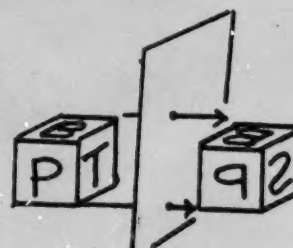


Figure 2.8. A reflection in 3-space.

In each space any rigid transformation is one of the three basic types or a product of two or more of them. Different orders of combining two transformations may result in different products.

Re-examination of the spatial tests in terms of the type of topological transformations involved illuminates the SR-0, Vz controversy. In SR-0 tests the transformations involved are rigid; in almost all

Vz tests other, nonrigid transformations are required although they may be combined with rigid transformations. For example, all the two-dimensional SR-0 tests (Flags, Figures, Card Rotations) require the subject to differentiate between rotations and reflections of a referent figure. All the variations of the Cubes test require rotation of cubes in three-space. In Spatial Orientation by Guilford and Zimmerman translation and rotation of a motor boat are required (but not reflection). The various form-board tests and Werdnig's Puzzles (all two-dimensional Vz) involve cutting a figure into parts or putting pieces together to form a larger figure; rigid transformations are often involved but usually only of the pieces. Most three-dimensional Vz tests are surface-development type tests which involve folding as does the two-dimensional Vz test Punched Holes.

Tests of visual disembedding usually require only a translation of the simple figure onto the complex stimulus. Perhaps this explains their loading on the SR-0 factor (French, Ekstrom, & Price, 1969). What is hard about these tasks is to disregard the unused line segments in the complex figure.

So far this mathematical analysis, although tidier than the psychological one, offers little more insight than does that of Michael, et al. Additional refinement is needed. Zimmerman (1954) has suggested that the SR-0 and Vz factors represent stages along a continuum of complexity or difficulty. He even hypothesized that by increasing item difficulty and complexity in one type of test it could be made to load on

Perceptual Speed, SR-0, Vz, and Reasoning in that order, and he managed to design three forms of one test which did indeed load on those first three factors in a male sample. Werdelin (1961) found that although items on two tests, Sections and Inversion, could be solved by a rigid transformation of the whole figure, this was too difficult for his tenth-grade subjects. Instead they broke each figure into parts and transformed them one at a time; these tests appeared on the Vz factor.

The difficulty of an item may have several sources. In general, the larger the number of degrees the item must be rotated, the harder the item. This author suspects that if none of the rotations in Cards, Flags, or Figures were more than 20° , these tests would measure Perceptual Speed. Visualizing transformations in three-space is probably harder than in two-space, especially if the stimuli are pictorial representations of three-dimensional objects. Nonrigid transformations may be harder to visualize than rigid ones. The results of topological transformations which are products of several basic ones are harder to imagine than results of the basic ones. For example, a form-board item requiring only that the pieces be slid together is probably easier than one in which some pieces must be rotated first. The fact that Spatial Visualization by Guilford and Zimmerman loads on Vz instead of SR-0 may be due to the fact that on some items the required transformation is the product of two rotations of the alarm clock.

It is this author's hypothesis that in solving spatial items the subject first tries to visualize the transformations of the whole stimulus because this method is fastest. When the visual manipulation of the whole becomes too difficult, the subject moves the figure in stages or breaks the stimulus into parts and moves the parts separately. When the item becomes even more difficult, the subject may try to code the visual information verbally or symbolically and reason logically about it (another type of mathematical transformation).

The complexity of an item may depend on the solver as well as on the item. This can be illustrated by examining methods of solving items on the Cubes test where the task is to decide if the two drawings are of the same or different cubes. For a problem similar to that in Figure 2.9 a subject may rotate one cube 90° around a line drawn through the front, upper, right corner and the back, lower, left corner (the hidden one); for these subjects this is an SR-0 item. Another subject may rotate the left-hand cube 90° around a line drawn through the

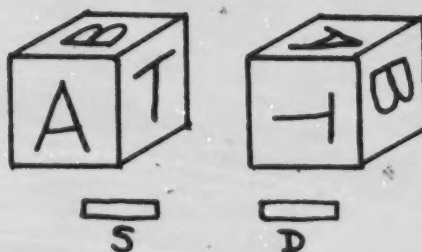


Figure 2.9. An item typical of the Cubes Test.

centers of the right and left faces (this puts the A on top) and then 90° around a line drawn through the centers of the top and bottom faces (this turns the A on its side). For these subjects the item is closer to a Vz item. Strategies such as these were reported by French (1965). The author's introspection of her own method of solving this item reveals a third, mostly logical method.

Alphabet letters in their usual position have a top and a bottom, a right and a left side. In both cubes the top of the A is next to the left side of the B. The bottom of the B is next to the top of the T. The right side of the A is next to the left side of the B in both cubes. Therefore they must be the same.

For the author these items probably measure the same figural reasoning ability as does a figure matrix test. Werdelin (1961) reported this method of solution of Cubes items, especially among girls. Probably the tests described in this chapter are more pure measures of the Space factor in populations with less well-developed reasoning ability and less training in geometry.

Thinking of spatial ability as the ability to visualize topological transformation of figures in a plane or in three-space enables one to explain sex differences both in mean performance and in factor structure. If it is true that more men than women transform space figures easily, then as a group they should perform better on most spatial tests because speed is more or less of a factor on all of them. Furthermore, if on a given test more women than men are forced to rely on logical manipulation of a symbolic coding of the figural material

(due to their more limited ability to visualize transformations) then the tests would have different loadings on spatial and reasoning factors for males and for females. These differences in factor structure have indeed been observed. The splitting of the Space factor into SR-0 and Vz observed only in male samples may be due to the fact that the range of items or tests on which many women can visualize the transformations may be too narrow to allow for the appearance of subfactors.

Highly capable girls with some knowledge of geometry (like those in Werdelin's sample) may be able to use that knowledge and their reasoning abilities to do almost all the tests as well as boys. This author suspects that the crucial sex-related difference in Werdelin's study was in ability to manipulate three-dimensional figures visually. Werdelin discounted dimensionality because one of the four three-dimensional tests did not show sex differences. However, that one (Cubes) permits logical solution by the method described in an earlier paragraph, and more time per item was allowed for this test than for Pairs of Cubes which showed sex-related differences significant at the .10 level. Items on Werdelin's Sections, a test on which the sex differences were most significant ($p < .001$), appear to this author to be solvable only by visualizing the transformation.

Causes of Sex-Related Differences in Spatial Abilities

One cannot end a chapter on sex-related differences in spatial abilities without some discussion of possible causes of those

differences. A number of different factors have been hypothesized to influence visual spatial performance, all of which have some credibility and none of which seems to provide the only answer. Since searching for antecedents is not the purpose of this study, this section of the review of research will be cursory rather than exhaustive.

The hypothesis that spatial ability is a trait determined by a sex-linked recessive gene on the X chromosome has been advanced by Stafford and supported by others. Stafford's (1961) theoretical prediction of parent-child correlations on the Identical Blocks test is supported by his data. On the basis of the same genetic hypothesis Bock and Kolakowski predicted that 50 percent of the males and only 25 percent of the females should have the trait. The results of fitting mathematical curves to the distributions of boys' and girls' scores on the Guilford-Zimmerman Spatial Visualization test in an eleventh-grade sample support the hypothesis. The lack of sex-related differences in some populations might be considered evidence to the contrary. However, the three groups in which no sex differences in spatial ability have been observed--Canadian Indians, Eskimos and Australian Aborigines--are racially different and possibly genetically different from the group in which genetic sources of the variance have been hypothesized.

Sex differences in socialization are also suspected of causing the sex-related differences in spatial skills. According to

Mitchelmore's (1975) survey of cross-cultural research, groups who hunt and wander usually have more highly developed spatial skills than those who farm. If boys in our culture wander more widely in their play than girls, this could be a cause of the sex differences observed in spatial ability. Another related finding of the cross-cultural research was that people in societies which encourage their members to be independent, objective, and self-sufficient had higher spatial ability than people in societies which valued the opposite characteristics. Mitchelmore even cited a generalization by Stewart that sex differences in spatial ability are lower or nonexistent in cultures where both sexes are brought up to be independent.

Spatial training may be the issue. Studies by Witkin and his followers have maintained that field-independence-dependence is a cognitive style rather than a cognitive ability or skill because performance does not improve with training. These conclusions have been criticized on the grounds that the training in those studies was not adequate (Kagan & Kogan, 1970; Sherman, 1967), and have been contradicted by Sherman's more recent work (1974). Other studies have shown that it is possible to improve spatial skills by training (Sherman, 1967; Wolfe, 1970). The fact that sex differences appear in some studies as early as fifth grade may indicate that sex differences in play activities or early school experiences are involved. Noting that girls do learn to read more easily than boys, Fennema (1974b) suggested that, because it is easy for them, girls may rely more

heavily on symbolic representations (or may be encouraged to) in early school learning. Perhaps this hampers development of spatial skills at school. If both at school and at play girls have less opportunity to learn and practice spatial skills, an early appearance of sex differences in this area should be expected. The Russians consider spatial training an important part of school learning and also report that their girls are at no disadvantage in spatial abilities (Chetverukhin, 1949/1971). However, no empirical research was cited.

Both hereditary and environmental causes of sex-related differences in visual spatial abilities are credible. The present state of the research does not enable one to decide between them or even to assign a portion of the variance to each. However, since socialization practices can be changed and spatial training improved whereas genetic differences are immutable, it seems more important to investigate the environmental causes of sex-related differences in visual skills.

Chapter III

REVIEW OF RESEARCH ON MATHEMATICAL PROBLEM SOLVING ABILITY

Introduction

The preceding chapter presented an explanation of sex-related differences observed in visual spatial abilities in terms of the topological transformations required by the tests. One might consider most spatial tasks to be problems in transformational geometry. Are women similarly handicapped in solving all mathematical problems? This chapter examines the evidence that sex-related differences in ability to solve mathematical problems do exist. It also investigates the hypothesis that the sex-related differences in spatial ability are at least partly responsible for differences in mathematical problem solving.

To qualify for inclusion in this review a study must use test items intended by their designers to measure mathematical abilities other than computation and test items which seem to this reviewer to satisfy the definition of a mathematical problem given in the introductory chapter. For example, studies of mathematical reasoning were often included. Selecting studies involved judgment because of the definition's requirement that the individual attempting to answer

the question or attain the objective not possess a memorized answer or an immediate procedure. The general policy was to include any doubtful studies.

Sex-Related Differences in Mathematical Problem Solving Ability

Reviews of sex-related differences in intellectual functioning have often noted differences in mathematical ability (Garai & Scheinfeld, 1968; Maccoby, 1966; Maccoby & Jacklin, 1974); a frequent generalization is that although girls may be better computers, boys excel at mathematical reasoning (Glennon & Callahan, 1968; Maccoby, 1966). In a detailed review devoted entirely to sex differences in mathematics learning Fennema (1974a) indicated that there is some support for this generalization at some grade levels but not at others. The purpose of this part of the chapter is to examine the evidence that sex-related differences exist in that ability which has been called quantitative reasoning or mathematical problem solving. The studies reviewed here are grouped in order to note when differences appear. The groupings are (1) elementary, (2) Grades 7 and 8, (3) Grades 9 through 12, and (4) college and adult.

Surveying the literature on sex-related differences in mathematical problem solving in primary school children is complicated by the fact that written tests are less common and those which are given seem to be mostly computational. For these reasons only studies of older elementary-school children will be discussed. A study done by

Parsley, Powell, and O'Connor (1964) using the California Arithmetic Test-Reasoning with subjects in Grades 4 through 8 found significant sex differences in favor of boys in 12 subgroups and in favor of girls in 7 subgroups out of a total of 75 comparisons. Despite the fact that no significant differences had been found in a previous study by the same authors (Parsley, Powell, O'Connor, & Deutsch, 1963) using the same test and that a difference of 5 out of 75 is barely above the .05 level of probability, the authors concluded that boys excel at arithmetic reasoning. In a study of sixth-grade students Jarvis (1964) found that boys at all ability levels surpassed girls in the area of arithmetic reasoning measured by the California Achievement Test Battery. In the CAA Project (Harris & Harris, 1973) two of the cognitive abilities tests contained mathematical problems: Number Series, given in 1970 and again in 1971, and Necessary Arithmetic Operations, given only in 1970. No significant sex-related differences were found in any of the means of their two fifth-grade samples.

The National Longitudinal Study of Mathematical Abilities (NLSMA) has been both the most intensive and the most extensive study in the field of mathematical abilities or achievement. Three different groups were used: the X-Population was tested in Grades 4 through 8, the Y-Population in Grades 7 through 11, and the Z-Population in Grades 10 through 12. The purpose of the NLSMA was to study the differences among users of various mathematics textbooks on a variety of measures of mathematics achievement. A content by level-of-behavior

matrix was used to categorize the mathematics scales. Number systems, geometry, and algebra were the three content areas. The four levels of behavior were computation, comprehension, application, and analysis. Definitions of the last two levels follow.

Application - Items designed to require (1) recall of relevant knowledge, (2) selection of appropriate operations, and (3) performance of operations. Items are of a routine nature. They require the subject to use concepts in a specific context and in a way he has presumably practiced.

Analysis - Items designed to require a nonroutine application of concepts. (Romberg & Wilson, 1969)

Clearly the latter category is the closer one to the definition of a mathematical problem in this study. However, since items from the application level may be nonroutine problems for some subjects, this review also discusses results on scales at the application level.

A word should be said about the statistical analyses in the NLSMA. First of all, sex-related differences were not the main interest in this study and were reported only for scales which were used in the textbook analysis and on which there was no sex by textbook interaction. The latter restriction excluded very few scales, none at the analysis level. However, the former excluded comparison within some groups at the high school level who did not elect college-preparatory mathematics courses and therefore were not using books under consideration. Secondly, the principal statistical method was analysis of covariance. The covariates used differed at each grade level but were usually, in this order, verbal IQ,

non-verbal IQ, and some combination of measures of mathematics achievement taken from NLSMA tests previously administered. Thus sex-related differences reported in the subsequent discussion are those remaining after removal of the variance due to IQ and mathematics achievement and significant at the .05 level.

Most of the scales used by the NLSMA in elementary school testing were at the computation or comprehension level. Two of the three applications scales given, both concerned with number systems, showed significant sex-related differences in favor of boys. There was only one analysis scale given in Grades 4, 5, or 6, a geometry scale, on which boys did significantly better than girls. (Carry & Weaver, 1969)

In another longitudinal study done by Hilton and Berglund (1974) subjects were measured in Grades 5, 7, 9, and 11 using the Sequential Test of Educational Progress - Mathematics (STEP-Math) and the School and College Ability Test, Quantitative Series (SCAT-Q). The authors considered both to be measures of mathematics achievement and regarded STEP-Math as a measure of ability to apply skills to problem solving in mathematics. The importance of using those tests, as the authors noted, is that on both STEP-Math and SCAT-Q the forms are vertically equated and the results can be treated as if the same test was taken at each level. The sample was divided into an academic group and a non-academic group according to what program they eventually pursued in high school, and results were analyzed separately by group. The fifth-

grade, non-academic girls did significantly better than the non-academic boys ($p < .05$) on the SCAT-Q; otherwise there was no significant difference.

A summary of the studies done at the upper elementary level shows that some sex-related differences in mathematical problem solving have begun to appear but the results are mixed. Boys' scores were better on the one NLSMA analysis scale, in the Jarvis study, and possibly in the second study by Parsley, et al. One group of girls did better on one test in the study by Hilton and Berglund. Otherwise, no significant differences appeared.

Results of problem-solving studies on subjects in Grades 7 and 8 have been grouped together. Although students in these grades may be ability-tracked in mathematics classes, usually the choice of the algebra-geometry sequence or the general mathematics sequence is made later. Hilton and Berglund reported a sex-related difference in favor of boys in the academic group on STEP-Math significant at the .05 level in seventh grade. No significant differences were reported for the academic group on SCAT-Q or for the non-academic group on either test. The NLSMA gave STEP-Math to the Y-Population in seventh grade, categorizing it as an application test, and also found boys' performance to be superior at the .05 level (McLeod & Kilpatrick, 1969).

Sex-related differences significant at the .05 level were also found in all but one of the analysis scales and the one application

scale designed by NLSMA for the X- and Y-Populations in seventh and eighth grade. In the X-Population the differences were significant on all analysis scales: two number systems scales given in Grade 7 and one given in Grade 8, one geometry scale given in Grade 7 and repeated in Grade 8, and a second geometry scale given only in Grade 7. On the one application (geometry) scale given to the X-Population in Grade 8, the difference was significant but the univariate F-value was smaller than for the analysis tests. (Carry, 1970) All of the analysis tests (three geometry, two number series, and one algebra) given to the Y-Population were given in Grade 8. Significant sex-related differences in favor of boys were reported for all but the algebra scale on which there was no significant difference. (McLeod & Kilpatrick, 1969)

In a study of problem-solving styles in high-ability, eighth-grade subjects Kilpatrick (1967) found that although the scores for boys and girls were about the same, girls used significantly more deduction ($p < .01$) and more equations ($p < .05$). Kilpatrick attributed this sex-related difference in the use of algebraic skills to the fact that the class in which some algebraic techniques were taught contributed only girls to his sample. In light of the NLSMA finding that the only scale on which boys did not excel was the one algebra scale, this conclusion might not be a valid one.

In the National Assessment of Educational Progress (NAEP) consumer math skills were measured by a test of problems given to 13-year-olds, 17-year-olds and young adults ages 26 to 35. A number of problems

were not given to the 13-year-old group because the content was presumed unfamiliar to them, but on the remaining 25 problems the boys' median was 1-1/2% above the median of the total group and the girls' median was 1-1/2% below. Boys did better than girls on all the percentage problems except that one requiring the averaging of wages. Females outperformed males by 12% on reading a pictograph on rural populations. However, in general the observed differences on individual items were small. (Ahmann, 1975)

With the exception of the NAEP all the studies discussed in this section in which sex-related differences were observed were conducted with students of above-average ability. The generalization is often made in NLSMA reports that the groups studied were above average in mental ability, mathematics achievement, and socio-economic status. In particular, the mean Y-Population performance on STEP-Math in seventh grade was in the 62 to 88 percentile band according to norms in the STEP manual (McLeod & Kilpatrick, 1969). In the Parsley, et al. study (1964), which was cited in the preceding section on elementary school subjects and which included seventh- and eighth-grade students, sex-related differences in arithmetic reasoning were most notable in the upper ability levels.

There is another indication that overall superiority of boys in mathematical problem solving at the seventh and eighth grades may be due to the superior performance of boys of high ability. In a study of mathematical precocity in students of this age (Stanley, Keating, &

Fox, 1974) the mathematics subtest of the Scholastic Aptitude Test (SAT-Q) was used as a screening device. Although the SAT-Q may not be a measure of mathematical problem solving of a nonroutine nature for high school seniors, the authors felt that the items on the test were nonroutine problems for these seventh and eighth graders who had had no formal training in high school mathematics. In the sample of students who volunteered for screening, boys' performance was superior to girls' at the .05 level at both grade levels, and the discrepancies increased with age.

Surveying the studies done on high school students requires additional caution. In many high schools, academic or college-preparatory and non-academic or vocational mathematics tracks are available; the proportions of boys and girls in those tracks may differ. Courses in mathematics also become elective at some point in most high schools. It is known that more boys than girls elect mathematics courses. For this reason, the fact that there are different norms for males and females on some college aptitude tests is not relevant to this discussion; boys in eleventh or twelfth grade as a group have had more opportunity to learn and practice mathematics than girls.

A good example of this lack of control for number of mathematics courses taken is the Project Talent survey discussed in the last chapter (Flanagan, et al., 1964). In the arithmetic reasoning test significant sex-related differences were observed at the ninth grade and

increased every year through twelfth grade. Although the content of problems was presumably taught before Grade 9, daily practice in mathematics skills at school must be useful in solving those items. Similar differences were observed on the intermediate and advanced mathematics tests. Likewise the NAEP which found a difference in male and female medians of about 6% at age 17 did not control for level or amount of mathematics taken. In the remaining studies of high school students reviewed here the students were all in the same class or track when tested.

The NLSMA populations were divided into two groups at each grade level. Students taking the usual college-preparatory course for that grade formed Group 1; others were in Group 2. Boys scored significantly higher than girls on both analysis scales (geometry and algebra) given to Group 1 of the Y-Population in ninth grade. No comparison was reported for the two analysis scales given to Group 2. (Kilpatrick & McLeod, 1971a) Group 1 of the tenth-grade Y-Population was given six analysis scales and four application scales, all with geometry content. Boys did better on two of the three applications scales. Girls did significantly better than boys on two analysis scales, Structure of Proof and Constrained Motion; boys did better on the remaining six tests. Again the comparison of boys' and girls' scores on the analysis tests given to Group 2 was not discussed. (McLeod & Kilpatrick, 1971) The tenth-grade Z-Population took the same two application measures

given to the Y-Population and STEP-Math. Boys did significantly better on all three. (Wilson & Begle, 1972)

Sex-related differences were investigated only in Group 1 of both the Y- and Z-Populations in eleventh grade. In the Z-Population boys excelled on the geometry analysis scale, but there was no significant difference on the algebra or number systems analysis scales. Two application scales were given (both geometry), one with a significant difference in favor of boys (Wilson, 1972). In the eleventh-grade Y-Population the same or similar analysis measures were used. In this sample boys did better on the number series and algebra scales, while there was no significant difference on the geometry scale. Four application measures were given, all geometry. On one there was a sex by textbook interaction, on the rest there was no significant difference (Kilpatrick & McLeod, 1971b). In the twelfth-grade report sex-related differences were not noted (Romberg & Wilson, 1972).

Taken as a whole the NLSMA results seem to provide overwhelming evidence of male superiority on mathematical problem solving tests (See Tables 3.1 and 3.2.) This impression should be tempered by several limitations of this study. First, it should be noted that since only the first group was used in the textbook analysis at each grade level, only the upper ability level was examined for sex-related differences. This restriction of the sample was more severe at the high school level than in Grades 7 and 8. The importance of making

Table 3.1

Sex-Related Differences in Applications Scales
National Longitudinal Study of Mathematical Abilities
X-, Y- and Z-Populations

Grade	Number Systems			Geometry			Algebra		
	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls
4	1	0	0						
5	1	1	0						
6	1	1	0						
7	1	1	0						
8	1	1	0						
9									
10				6	4	0	1	1	0
11				6	1	0			

this distinction is supported by examination of the Hilton-Berglund (1974) data for Grades 9 and 11. In the academic group boys' scores were significantly higher ($p < .001$) on STEP-Math in Grades 9 and 11

Table 3.2

Sex-Related Differences in Analysis Scales
National Longitudinal Study of Mathematical Abilities
X-, Y- and Z-Populations

Grade	Number Systems			Geometry			Algebra		
	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls
4									
5				1	1	0			
6									
7	2	2	0	2	2	0			
8	3	3	0	4	4	0	1	0	0
9				1	1	0	1	1	0
10				6	4	2			
11	2	1	0	2	2	0	2	1	0

and on SCAT-Q in Grade 11. In the non-academic group the boys' scores were significantly higher only in Grade 11 and the difference on SCAT-Q barely achieved the .05 level of significance.

Another question that can be raised about the NLSMA results involves the use of analysis of covariance. While trying to describe a factor which appeared in several of his studies and which he called Mathematical Reasoning, Werdelin (1958) noted that the tests loading on this factor had verbal, computational or numerical, logical, and visual-perceptual elements in varying amounts. If in the NLSMA the variance due to verbal IQ, non-verbal IQ, and mathematics achievement was removed before the sex comparisons were made, then the sex-related differences observed are only in what remains. Changes which can be produced by this kind of statistical removal can be seen in a study of problem solving in ninth-grade algebra done by Sheehan (1968). Although girls had a slight (but nonsignificant) advantage on the problem-solving scores, by statistically removing the variance due to eighth-grade mathematics achievement, algebra aptitude, and preinstructional knowledge of algebra, Sheehan was able to produce a significant sex difference in favor of boys. He concluded that boys learn to solve algebra problems better than girls. There is no way to tell from the NLSMA Reports how much of the variance in the analysis scales remained after removal of the covariates; it may have been a small part. One can, however, get an estimate of the proportion of the variance due

to sex by using the ω^2 statistic (Hays, 1973). Application of ω^2 to the three analysis scales given to the eleventh grade Y-Population shows that less than 1% of the variance was due to sex.

Not only might the significant sex-related differences exist on only a small part of the analysis component, that part might consist predominantly of the visual-perceptual elements. That sex differences do exist in this area was discussed in the last chapter. Looking at sex-related differences on the NLSMA analysis scales for Grades 7 through 11 separately by content suggests that the spatial component accounts for a portion of those differences. On the number series scales the items were practical problems in which visualizing the real situation might be helpful; boys excelled on 86% of these scales. They excelled on 87% of the geometry scales but only 50% of the algebra scales. One of the geometry scales on which girls excelled was Structure of Proof, which seemed to require no spatial skill.

Another factor related to problem content may have produced some of the sex-related difference in the number series scales. These were the only problems about people. In all cases in which the sex of the person was specified, the people were male. The influence of this sex stereotyping on the scores cannot now be measured but it cannot be overlooked.

In Werdelin's (1961) study of highly capable tenth-grade students several tests could be identified as tests of mathematical problem solving: Geometrical Problems I and II, Arithmetic, Number Series and

probably Geometrical Construction. In the original samples boys excelled girls on Geometrical Problems I and II and on Arithmetic at least at the .01 level; boys' performance on Geometrical Constructions was significantly better than girls' at the .05 level. In a subsample of boy-girl pairs matched on age, numerical ability, and general reasoning ability sex-related differences remained on Geometrical Problems II ($p < .001$), Geometrical Problems I ($p < .01$), and possibly Geometrical Construction ($p < .10$). In order to identify the source of the differences Werdelin investigated the sex-related differences on individual items on these three tests. Werdelin concluded after this analysis that the sex difference was essentially in the ability to visualize the material and reorganize it. After examining the problems, this author would categorize the problems showing significant sex differences as those requiring the visualization of several successive geometric transformations, especially in three-dimensional space.

The caution exercised in the review of the high school studies must be repeated when surveying research done on college students and other adults. Very (1967) found sex-related differences significant at the .01 level on virtually all of his tests of mathematical problem solving. However, the results are questionable due to lack of control for mathematics background. Likewise the results of the NAEP (Ahmann, 1975), which showed a sex difference in solving consumer problems of about 13% in young adults, should be evaluated critically because of this lack of control.

The most significant body of research on mathematical problem solving in college students is a group of interrelated studies done first at Stanford and then at Yale. Although the researchers were psychologists and their intent was to study human problem solving in general, their measures consist almost exclusively of problems which satisfy the definition given in the introduction. After Sweeney's 1953 study in which sex-related differences in addition to those due to intellectual factors were found, the others investigated various other noncognitive sources of the difference.

Sweeney conducted several different studies using three different sets of problems. The tool problems (Set A) are highly related to mechanical knowledge and not really mathematical; they are not discussed further. Sweeney divided Set B (puzzle book problems) into three types: the three verbal posers require changing the use or function of something in the problem; three of the four figure problems require dividing the figure into pieces of the same size and shape, while one was a match problem; the three extension problems require the subject to continue a line of reasoning already begun. Sweeney felt that the first two types required "restructuring" whereas the last one did not. In a group of men and women matched on a verbal IQ measure sex-related differences in favor of males were reported only for verbal posers ($p = .006$) and figure problems ($p = .07$).

Sweeney's mathematics-logic problems (Set C) were also divided into three types: in the two caution problems the obvious answer is wrong:

the five glare-and-guess problems were supposed by Sweeney to require less "restructuring" than the five ingenuity problems. In this author's opinion the glare-and-guess problems are typical of those solved in first courses in algebra whereas the ingenuity problems are not. Sex-related differences were found in caution ($p = .005$), glare-and-guess ($p = .003$), and ingenuity ($p = .0001$) problems even after matching on IQ. Differences on the first two types were eliminated by matching on computation, years of mathematics taken, mathematics information, or score on the Flags test, one variable at a time. Sex-related differences on the third type were reduced but did not disappear after each of these four matchings. In still another study he found that matching on the SAT-Q did not remove the sex-related differences in problem solving.

Sweeney's study can be criticized on several points. One is the small number of problems of each type used. Also Sweeney himself acknowledged that the matching technique was weak. The study indicated that a number of variables were important to solving mathematics problems, but due to the size of his sample he could not match on more than one variable at a time. On the other hand, his work led others to follow with studies examining these factors one at a time or using more powerful statistical techniques.

Carey (1955) investigated the role of attitudes toward problem solving using problems from Sweeney's study on which sex-related differences had been observed. She found that men's attitudes toward problem

solving were more favorable than women's, that their problem-solving performance was higher, and that there was significant correlation between the two. More importantly, following a treatment designed to improve attitude, women's problem-solving performance improved significantly, whereas men's did not. This supports the hypothesis that women's poorer attitudes are at least partly responsible for their poorer performance in solving mathematics problems.

Milton (1957, 1958) and P. C. Berry (1958, 1959) investigated the relationship between sex role and problem solving using the Terman-Miles masculinity-femininity index and problems similar to those used by Sweeney and Carey. In Milton's first study problem-solving skill was significantly related to masculinity in both college men and women even after the effects due to verbal and quantitative SAT scores had been statistically removed. In a replication with high school students which controlled for IQ the correlation was significant only for women. In both of Berry's studies the correlation between masculinity and problem solving was significant only before statistical removal of the effects of both SAT scores. The validity of masculinity-femininity scales has been questioned by others. The scales were not particularly useful in these studies.

Actually Berry had hoped for much more. In the 1958 study he gave each subject the Terman-Miles and the Franck masculinity-femininity indices, two forms of Witkin's Rod and Frame test, Concealed Figures, Carey's attitude scale, a similar, supposedly more sensitive, attitude

scale by Block, and the Sarason Test Anxiety Scale. As he expected, there were significant sex-related differences on all of these tests and on the problem test, and most of these scales were significantly correlated with the problem test. What he had hoped was that each of these correlations would remain significant when the effects of the verbal and quantitative SAT's had been statistically removed. (That was a big order since the SAT-Q accounted for 50% of the variance.) He also hoped that whatever variance these nonintellectual variables had in common with problem solving, but not with verbal and quantitative ability, they would have in common with each other. This residue was going to be called sex role by Berry. The data did not support these hypotheses. After removing the effects of both SAT's, none of the other variables correlated significantly with the problem score for women. For men in both samples the tests remaining significant were the Rod and Frame, Concealed Figures, and Carey attitude tests. It should be noted that at the time of his study the spatial nature of the Rod and Frame and Concealed Figures tests was not widely recognized.

In another attempt to determine whether or not sex-related differences in problem-solving performance was related to "sex-role congruence," Milton tested men and women with pairs of mathematically equivalent problems, one "feminine" in content and the other "masculine." (Needless to say the sex-role stereotyping was incredible.) On each test ten of each type were used, but the combinations differed. As

usual, the men solved significantly more problems on the total test. However, this difference was reduced to nonsignificance on the feminine subtest. Hoffman and Maier (1966) also investigated this area. Although they found no significant differences, this cannot really be considered a failure to replicate. Their study eliminated all the geometry problems, gave each subject an entirely masculine or entirely feminine test and confounded the results with a motivational pep talk. Although the NAEP (Ahmann, 1975) noted superior male performance on such "feminine" problems as computing the time required to roast a turkey, mathematics background may have been the real determinant.

Summarizing the research on sex-related differences in mathematical problem solving is difficult. As was the case with visual spatial abilities the differences may be small, but they do seem to be there, even after controlling for mathematics background. Like spatial ability the differences seem to appear with early adolescence and may increase with age until maturity. The studies reviewed in this section indicate that the sex-related differences may be limited to the upper-ability level and to problems whose content is spatial or sex-biased. What little research there is on sex bias in problem content has already been discussed. In the next section the relationship between spatial abilities and mathematical problem solving is explored.

The Relationship Between Mathematical Problem Solving
Ability and Visual Spatial Abilities

The fact that sex-related differences in both visual spatial abilities and mathematical problem solving ability begin to appear in the upper elementary grades and develop throughout adolescence suggests that there is some relationship between them (Fennema, 1974b). The last section showed that both in Werdelin's work (1958, 1961) and in the NLSMA sex-related differences were more pronounced in tests whose mathematical problems appeared to be spatial. The classification of spatial tests in terms of the topological transformations they involve and the possible use of mathematical reasoning in spatial tests discussed in Chapter II also indicate that spatial and mathematical abilities overlap.

Smith (1964) has given evidence of the importance of spatial ability in his discussion of studies (mostly factor analytic) of mathematical ability. He said that spatial ability appeared to be important in tests which are genuinely mathematical as distinct from those which are merely mechanical or computational. That spatial skills become more important as the cognitive complexity of the mathematical task increases was demonstrated in a small-scale study by Anglin, Meyer, and Wheeler (1975). Both fourth- and sixth-grade students were given Form Board 1 and Card Rotations from the NLSMA battery. The mathematics achievement measures, which had been given the year before, had computation and concepts components for the fourth grade, and

computation, concepts, and applications components for the sixth-grade. In the sixth-grade results, the correlation coefficients of both spatial tests with the mathematics achievement measures increased in order of complexity; computation, concepts, and applications. The authors suggested that the lack of significant relationship in the fourth-grade results was due to the fact that their achievement measures did not involve higher cognitive processes to as high a degree as the sixth-grade measures.

In this section's review of the relationship between mathematical problem solving and visual spatial abilities, the results of each study are examined with these questions in mind.

1. What is the evidence of a relationship between mathematical problem solving and visual spatial abilities?
2. Is the relationship different for males and for females?
3. If more than one measure of spatial skills was given, are the relationships with the test of mathematical problem solving different for different spatial tests?
4. If more than one problem solving measure was used, are they differently related to the test(s) of visual spatial abilities?

In order to examine the hypothesis that what tests of spatial ability and mathematical problem solving have in common is figural reasoning ability (or in Guilford's terms, cognition of figural relations), comparisons with those tests will be made whenever the data are available. As is the case in the previous sections, the discussion is organized by grade level.

In the upper elementary grades the CAA Project (Harris & Harris, 1973) provided information on the relationships among spatial ability, figural reasoning, and mathematical problem solving. In the 1970 study of the CAA Project there were two space tests (Space Relations which is a type of form board test and Card Rotations) and two tests of mathematical problem solving (Number Series and Necessary Arithmetic Operations). Comparison of both was made with Figure Matrix and Figure Analogies, tests of figural reasoning. Table 3.3 presents the correlations coefficients among these tests.

Table 3.3

Correlations Among Three Abilities in the CAA Project (1970)
Mathematical Problem Solving, Figural Reasoning, and
Spatial Ability

	Card Rotations		Space Relations		Figure Analogies		Figure Matrix	
	Boys ^a	Girls ^b	Boys	Girls	Boys	Girls	Boys	Girls
Necessary Arithmetic Operations	37 ^c	44	42	41	59	42	58	59
Number Series	36	45	45	45	64	65	60	56
Figure Analogies	37	48	46	53				
Figure Matrix	40	51	52	47				

^a_n = 172

^b_n = 210

^c decimal points omitted

Note: Data from Harris and Harris, 1973

Although all of the correlation coefficients in Table 3.3 are significantly different from zero at the .01 level, none of the sex-related differences are significant even at the .05 level. Nevertheless, several trends are indicated which show up in studies discussed later. Although only for Card Rotations, the relationships with the problem-solving tests and figural reasoning tests in the CAA study were more significant for girls than for boys. Both the spatial tests and the mathematical problems tests were more closely related to the figural reasoning tests than they were to each other, which suggests that figural reasoning may be a bridge between the two. At least for boys Card Rotations was less related to the other tests than was Space Relations. Anglin, et al. (1975) found that the relationships of their mathematics achievement measures to Form Board 1 were more significant than those to Card Rotations for the sixth-grade group (boys and girls together). All the relationships with Form Board 1 were significant at the .05 level whereas for Card Rotations only the relationship with the applications scale was significant at the .05 level.

In the 1971 CAA study only Space Relations, Number Series, and Figure Matrix were used (see Table 3.4). As in the 1970 results, Figure Matrix seems to span the distance between the other two. Although there was no sex-related difference between the correlations for Number Series with Figure Matrix, there was a sex difference between the correlations of Number Series with Space Relations (significant at

the .08 level). There was a similar, although smaller, difference between the correlations of Figure Matrix and Space Relations for boys and for girls. It is interesting to note that in the 1970 study where the correlations of all the other tests with Space Relations were quite similar for boys and girls, there was no noticeable difference in the mean performance of boys and girls on Space Relations. In the 1971 results where the sex-related difference in performance on Space Relations was significant at the .05 level, the correlation pattern for that test was different. This lends support to the hypothesis discussed in Chapter II that sex-related differences in visual spatial abilities occur when girls as a group solve spatial items more often by reasoning methods (which may be slower) than by visualizing the transformations.

Table 3.4

Correlations Among Three Abilities in the CAA Project (1971)
Mathematical Problem Solving, Figural Reasoning, and Spatial Ability

	Space Relations		Figure Matrix	
	Boys ^a	Girls ^a	Boys	Girls
Number Series	37 ^b	51	61	60
Figure Matrix	43	53		

^a_n = 204 ^b decimal points omitted

Note: Data from Harris and Harris, 1973

With the NLSMA data there are two ways to investigate the relationship between the spatial tests and the analysis or application measures; one involves correlation coefficients, and the other involves analysis of variance. NLSMA Report #33 (Wilson & Begle, 1972b) detailed the intercorrelation of mathematical and psychological variables which included the spatial tests Card Rotations, Hidden Figures, Form Board, and Paper Folding, which is similar to Thurstone's Punched Holes and is a two-dimensional V_z test. Some of the populations took some of the spatial tests more than once which was indicated in the reports by giving them different numbers. Two forms of Hidden Figures and Form Board were used; in each case Form 1 is easier than Form 2. Unfortunately no correlations were reported for either one of the Form Board tests. Table 3.5 presents correlations of the analysis scales with the spatial tests for Group 1.

Several things should be noted about the correlations reported in Table 3.5. Since the smallest sample size reported for any correlation matrix was 500, a correlation coefficient of .11 was statistically significant at the .01 level. However, a correlation between two tests of .11 means that about 1% of their variance is shared. Despite the generally low relationships and the probability that none of the differences in this study were statistically significant, some interesting patterns can be observed. Paper Folding which requires a number of successive nonrigid transformations was a better predictor of all the analysis scales than either Card Rotations or Hidden Figures which

Table 3.5

Correlation Coefficients of Spatial Tests and Analysis Scales
in the NLSMA

Grade	Number Systems	Geom.	Alg.	Card Rotations			Hidden Figures				Paper Folding PY 609
				PY 220	PY 608	PZ 220	(1) PX 218	(2) PX 820	(2) PY 223	(2) PZ 223	
5 S ^a		X308					46				
7 F ^b	X601						41				
7 F	X611						40				
7 F		X603					38				
7 S		X715					45	23			
8 S	X911						33	11			
8 S	Y309			19	17				14		25
8 S	Y310			11	13				16		23
8 S		X912					41	21			
8 S		Y314		18	17				18		27
8 S		Y315		22	22				22		37
8 S		Y316		25	25				21		34
8 S			Y312	18	18				21		25
9 S		Y506		15	19				17		34
9 S			Y507	18	22				20		36

^aSpring

^bFall

^cdecimal points omitted

Note: Data from Wilson & Begle, 1972 (#33)

Table 3.5 (Continued)

Correlation Coefficients of Spatial Tests and the Analysis Scales
in the NSLMA

Grade	Number Systems	Geom.	Alg.	Card Rotations			Hidden Figures				Paper Folding
				PY	PY	PZ	(1) PX	(2) PX	(2) PY	(2) PZ	PY
				220	608	220	218	820	223	223	609
10 F	Y601			20	24				20		34
10 F		Y602		19	25				21		38
10 S		Y702		14	16				15		28
10 S		Y703		01	09				08		19
10 S		Y704		22	26				18		39
10 S		Y705		16	18				16		35
11 F		Y806		08	12				13		29
11 F		Y807		01	08				12		21
11 S	Y910			16	12				14		19
11 S	Z310					16				13	
11 S		Y909		19	17				19		27
11 S		Z303				18				14	
11 S			Y911	11	08				13		11
11 S			Z306			15				10	

^a Spring

^b Fall

^c decimal points omitted

Note: Data from Wilson & Begle, 1972 (#33)

require one rigid transformation. At each grade level the spatial tests were more highly related to the geometry scales than to the algebra or number systems scales. The correlation coefficients also decreased with age. This was probably due to the fact that mathematical problem solving at the upper levels appears to require more specific mathematical knowledge.

Selection of correlation coefficients out of a larger matrix always leads to some distortion. Although the relationships of the other mathematics scales to the spatial tests are not presented in Table 3.5, they were investigated by the author. In all but a few isolated cases the correlations of other mathematics scales in the NLSMA with the spatial tests were no higher than those of the analysis tests. Other tests whose correlation coefficients with the space tests were sometimes as high were the various forms of STEP-Math, geometry comprehension or applications scales, the Lorge-Thorndike Non-Verbal IQ scales, and some number systems scales on fractions which presented figural stimuli. Another limitation is that only the results of Group 1 are presented in the table. However, examination of the correlation matrices for Group 2 (a group which took fewer analysis measures, especially in geometry) reveals the same patterns although generally less significant correlations.

Additional information on the relationships between the analysis or applications scales and the spatial tests in the NLSMA was generated

by a two-way analysis of variance (Wilson & Begle, 1972a). Analyses were done separately by sex for each pair of tests. Table 3.6 presents the percentages of the total number of analysis or application scales of each content type on which significant main effects ($p < .05$) of

Table 3.6

Per Cents of Total NLSMA Analysis and Application
Scales on Which Significant Main Effects
of Spatial Variables Were Observed

	Applications		Analysis		
	Number Systems	Geom- etry	Number Systems	Geom- etry	Algebra
Card Rotations					
Boys	20	36	30	27	13
Girls	10	29	20	35	13
Paper Folding					
Boys	70	86	50	81	50
Girls	80	57	90	88	25
Hidden Figures					
Boys	70	50	13	44	33
Girls	40	70	0	33	0
Form Board					
Boys	50	67	37	33	17
Girls	60	50	37	58	0

Note: Data from Wilson and Begle, 1972a (#26).

spatial variables were found. Again Paper Folding seems to be the most closely related of all the spatial tests to both the analysis and the application measures. Form Board was also significantly related to a higher percentage of the analysis and application scales than Card Rotations, although not than Hidden Figures. This lends support to the hypothesis that Vz tests have more in common with mathematical problem solving than SR-O tests. As with the correlation coefficients, spatial tests were more closely related to the geometry scales in this study than to the algebra or number systems scales. There appears to be no pattern to the differences in the percentages for boys and for girls.

Dodson (1972) used a discriminant analysis of the NLSMA data to determine the characteristics of successful, insightful problem solvers. From the analysis scales given to the eleventh-grade Z-Population, he selected 14 geometry problems, 19 algebra problems, and 7 number problems which he and two others judged to be nonroutine. Dodson then divided each of the subtests and the total test into three to six levels of performance and used discriminant analysis to see which other tests given by NLSMA discriminated among those levels. He chose .001 as his level of significance but reported discriminations at the .01 and .05 levels. Hidden Figures 2 discriminated at the .001 level for the total test and the geometry and number subtests; it was not related to the algebra subscale. Card Rotations discriminated at the .001 level for the total test and the geometry subtest but only at the .01 level for the algebra and number subscales. Paper Folding and Form Board 2 were

not included in the analysis, and there was no separate analysis by sex. The variables which were better discriminators than these two spatial ability tests were all the mathematical scales and most of the cognitive reasoning scales including Pictorial Analogies.

In Werdelin's 1958 study the interrelationship of mathematical problem solving and visual spatial abilities in male subjects was demonstrated by factor analyses of several separate sets of data. The Space factor included those of his tests which are usually considered spatial, including Gottschaldt I and II, and two tests of mathematical problems, Number Series I and II. The Mathematical Reasoning factor included four tests of mathematical problems and all the Vz tests, both two- and three-dimensional, although their loadings were lower than those of the mathematics measures. Werdelin's later study of geometry and the space factor (1961) used both male and female subjects. Cubes had a low but significant loading on the General Reasoning factor which included all the tests with mathematical problems. All four of the geometry tests and Number Series had significant loadings on the Space factor.

Data from the Project Talent study (Flanagan, et al., 1964) also make it possible to compare results on different spatial tests separately for boys and girls. Comparisons were reported only for the twelfth-grade sample and are summarized in Table 3.7. Two-Dimensional Space is an SR-0 test like Thurstone's Figures; Three-Dimensional Space is a surface-development type of test. Mathematics I is a test of

arithmetic reasoning; Mathematics II is primarily algebra with some intuitive geometry; the content of Mathematics III is a plane and solid geometry, algebra of quadratic equations, trigonometry and analytic geometry. Abstract Reasoning is a figure matrix test like that used in the CAA study. As has been noted earlier, selecting a submatrix of a correlation matrix introduces some distortion. In the complete matrix there are no tests whose correlations with the spatial tests are any higher than those of the mathematics measures. Others in the same neighborhood are tests of information in mathematics, physical science, electronics or mechanics, the mechanical reasoning test, and the creativity test.

Table 3.7

Relationships Among Three Abilities in the Project
Talent Study Mathematical Problem Solving,
Figural Reasoning, and Spatial Ability

	2-Dimensional Space		3-Dimensional Space		Abstract Reasoning
	Boys	Girls	Boys	Girls	Boys
Mathematics I	32	32	46	48	53
Mathematics II	30	30	42	44	53
Mathematics III	28	23	37	35	37
Abstract Reasoning	40	37	54	52	

Note: Data from Flanagan, et al., 1964.

The relationships discernible in Table 3.7 follow some of the same patterns observed in other studies. The strength of the relationships decreased as the difficulty of the mathematics tests increased. The Vz test accounted for more of the variance on each mathematics test than the SR-O test although the dimensional variable may have been influential. There appear to be no significant sex-related differences in the coefficients. Again the figural reasoning test was related more strongly to both the spatial tests and the mathematical problems tests than the latter two were to each other. As an interesting aside, in a series of aptitude-treatment interaction studies using graphical and analytic modes of teaching quadratic inequalities (Carry, 1970; Webb & Carry, 1975; Eastman & Carry, 1975) the expected interaction did not occur until the authors replaced Paper Folding with a figure matrix test as a measure of spatial visualization.

The possibility that it is figural reasoning which visual spatial abilities and mathematical problem solving have in common is also supported by a study of intermediate algebra students done by Kline (1959). The two spatial tests he used did not form a separate factor but appeared with about equal weights on two factors. One, which Kline labeled Deductive Reasoning, included all of what appear to be the algebra problem tests. The other, which he called Adaptability to a New Task, had no tests with algebra tasks on it. In French's (1965) factor analytic study Figure Analogies always appeared on the Space Factor.

Some information on the relationship between visual spatial abilities and mathematical problem solving in college students can be obtained from the studies by Sweeney (1953) and P. C. Berry (1958, 1959) discussed in the previous section of this chapter. Sweeney found that matching on performance on the Flags test was as effective as matching on years of mathematics taken in removing or reducing sex-related differences in performance on his problem solving tests. In Berry's 1958 study of Swarthmore undergraduates and his 1959 study of Yale undergraduates (all men) the problem solving test was most highly correlated with SAT-Q for both men and women which is to be expected. What is surprising is that for that sample of Swarthmore men Concealed Figures was almost as highly correlated with mathematical problem solving as was the SAT-Q. For the samples of Swarthmore women and Yale men the correlation of Concealed Figures with the problem measure was third highest, following only SAT-Q and Carey's attitude scale. The correlation with the verbal SAT was the fourth highest for both sexes in both studies. Table 3.8 summarizes the relationships in both Berry studies. All the correlation coefficients were significant at the .05 level except that of Concealed Figures with problem solving after the removal of SAT-Q for women. In the Swarthmore study the correlations of Concealed Figures with the visual subtest were slightly higher than those with the mathematical subtest but not significantly so. Subtest correlations were not reported in the Yale study.

Table 3.8

Correlation Coefficients with Concealed Figures and
Problem Solving in Students at Swarthmore and Yale

	Swarthmore		Yale
	Men ^a	Women ^b	Men ^c
Problem Solving Total	68 ^d	40	35
Problem Solving Total (after removal of both SATs)	55	20	27
Subtests (no SAT removal)			
Mathematical	54	31	
Visual	68	33	
Straightforward	65	34	
Restructuring	58	33	

^a n = 55 ^b n = 48 ^c n = 79 ^d decimal points omitted

Note: Data from P. C. Berry (1958, 1959)

A summary of these correlational studies answers some of the questions posed at the beginning of this section. Visual spatial abilities appear to account for part of the variance in mathematical problem solving, but that part appears to decrease as the subjects get older or the mathematics measure depends more on specific mathematical information. In most cases the correlations of the spatial tests with the problem-

solving measures, although small, are at least as large as the correlations with any other test. The one exception occurs when tests of figural reasoning are included; these tests are more closely related to both spatial and problem-solving tests than the latter two are to each other. Spatial tests are related to problems with different content increasingly in this order: geometry; practical situations or arithmetic; algebra. Tests of the Vz factor are more significantly related to mathematical problem solving than are SR-O tests; tests of visual disembedding may fall somewhere in between.

The question of whether or not there are sex-related differences in the relationships is harder to answer. In the CAA Project it seemed that visual spatial ability and mathematical problem solving were more closely related for girls than for boys. This was not borne out in the two high school studies, NLSMA and Project Talent. However, the NLSMA data is voluminous and in an attempt to organize it, the reviewer may have obscured differences that really existed. On the other hand, age may be an important variable. The fact that the relationship of Concealed Figures to mathematical problem solving was stronger for men than women in the Swarthmore study should not be taken too seriously. In the replication with Yale students the correlation coefficients were about the same as for the Swarthmore women. In another study of high-ability college students (Bieri, Bradburn, & Gaiinsky, 1958) the correlation of Embedded Figures with SAT-Q was higher for women (.50) than for men (.40).

The trouble with evidence gathered from correlational studies is that the cause and effect relationships remain hidden. Does one use spatial skills in solving mathematical problems or mathematical skills in solving spatial items? The NLSMA found that a year's training in geometry did not increase scores on Card Rotations (Romberg & Wilson, 1969). If the NLSMA had asked the question about Paper Folding or Form Board 2 or if the geometry courses had been transformational, the answer might have been different. Nevertheless, it is usually assumed in the literature that spatial ability is somehow more fundamental than the ability to solve mathematical problems which involves other components as well. The question which arises next is how are spatial skills used to solve problems in mathematics. To answer that, one has to turn to introspective as well as experimental research.

The introspective accounts by mathematicians of their own problem solving or theorem proving are rich with visual imagery. In his classic work, The Psychology of Invention in the Mathematical Field, Hadamard (1945) described the role that visual imagery played in his thinking.

Now, personally, if I had to think of any syllogism, I should not think of it in terms of words--words would hardly allow me to see whether the syllogism would be right or wrong--but with a representation analogous to Euler's [a Venn diagram], only not using circles, but spots of an undefined form, no precise shape being necessary for me to think of spots lying inside or outside each other [p. 76].

He then described the visual images he employed in reconstructing the proof of the theorem that the sequel of prime numbers is infinite, and commented on the use he made of the image.

I need it in order to have a simultaneous view of all elements of the argument, to hold them together, to make a whole of them.... It does not inform me on any link of the argument (i.e., on any property of divisibility or primes); but it reminds me how these links are to be brought together [p. 77].

Hadamard noted other mathematicians whose mathematical thought was highly visual, especially in the intuitive or inventive stage. A letter he received from Einstein is particularly instructive, especially on the relative importance of words and images (in Einstein's case visual images).

The psychical entities which seem to serve as elements in thought are certain signs and more or less clear images which can be "voluntarily" reproduced and combined.

... But taken from a psychological viewpoint, this combinatory play seems to be the essential feature in productive thought--before there is any connection with logical construction in words or other kinds of signs which can be communicated to others [p. 142].

Taken together these two accounts suggest that the role imagery plays is to record the relationships or patterns among the elements of the problem and to facilitate the combination of elements into new patterns. The latter seems to be a transformation of the visual images involved.

However, Hadamard acknowledged that there are individual differences in style of mathematical thinking. Poincaré (1929) noted that

this difference has nothing to do with content as there have been geometers whose thinking was logical and symbolic as well as algebraists whose thought was full of visual imagery. This is true in ordinary people as well as in creative mathematicians. Menchinskaya (1946/1969) used subjects' introspection or retrospection to explore the nature of intellectual activity in solving arithmetic problems. Some people reported total reliance on verbal or numerical symbols. On the other hand, several highly visual methods were reported even for problems such as this:

The difference of two numbers is 50 and one number is twice as great as the other. What are the numbers?

The subjects using visual methods used the images to record the relationships in the problem and to transform the elements into new patterns.

Thus far the discussion has been concerned with the use of the mental images. Other aids to mathematical problem solving are diagrams--the visual images externalized on paper, chalkboard, or media screen. The role of the diagram in recording relationships in geometry appears to be very important. Sherrill (1970) constructed three forms of a problem-solving test in tenth-grade geometry. Although the verbal presentations of the problems were the same, the problems on the first form were accompanied by correct diagrams, on the second by no diagrams, and on the third by incorrect diagrams. Significant differences were found between all pairs in the above order. A replication of the

study with preservice elementary school teachers produced the same results even though these subjects were advised that the diagrams could be wrong (Webb & Sherrill, 1972). Botsmanova (1960/1972) noted that elementary-school children solved arithmetic problems accompanied by pictures illustrating the mathematical structure of the problem faster than problems accompanied by pictures merely illustrating the concrete situation.

When doing geometrical research Hadamard noted that he usually had a mental view of the diagram itself although usually an incomplete one. He reported that he often abstracted a part of the diagram and considered it apart from the rest. This suggests that visual disembedding is an important skill in solving geometry problems. A visual disembedding task was one of the measures Yakimanskaya (1959/1970) used to compare students using primarily verbal or primarily visual methods in solving geometry problems. The verbal group relied almost exclusively on the verbal condition and its implications and had difficulty focusing on parts of the diagram not discussed in the verbal condition. They reported solving the disembedding task by noting that the simple figure was two triangles with a common side and then searched for the common side in the complex figure. The visual group relied heavily on the diagrams in solving problems and often had difficulty seeing the implications of the verbal conditions. This group recognized the simple figure embedded in the more complex ones by a purely visual comparison.

Kabanova-Meller (1950/1970) noted three modes of solving geometry problems accompanied by diagrams; the modes varied in the use the subject made of the diagram. There were strong students who preferred each one of the modes. For weak students, difficulty in using diagrams arose not only in disembedding simple figures from complex drawings, but also in foreseeing which simple figures would be useful. Kabanova-Meller attributed this weakness in the subjects to the fact that when they had learned theorems they had not generalized them to transformations of the diagrams accompanying the theorems. She devised a treatment which was successful in overcoming this weakness. Wolfe (1970) was able to improve performance on problems accompanied by diagrams by a spatial visualization training program.

Being able to pick out a simple figure from a complex diagram and being able to recognize an element of a diagram as the transformed image of the diagram accompanying a learned theorem are both important to solving mathematical problems. One area not yet discussed is the ability to represent the information in a problem by a diagram.

Botsmanova (1960/1972) trained a group of elementary-school students to draw diagrams when solving arithmetic problems. On the criterion measure the experimental group drew more diagrams and solved more problems with fewer incorrect trials than either the control group which had the same amount of training in problem solving but without diagrams or the control group which had the regular school program. Kilpatrick (1967)

noted that drawing a diagram was not significantly related to success in problem solving in his eighth-grade sample. He did not note whether or not the diagrams drawn were correct representations of the problems.

Sex-related differences are not reported in any of the Russian studies, (probably due to their political philosophy) and Hadamard seems to have missed the opportunity to examine the mathematical thought of Emmy Noether. However, there is some evidence that girls perform less well at drawing diagrams. Mitchelmore (1975) found that fifth-grade Jamaican girls were less competent at drawing three-dimensional solids from models than boys; this surprised him because in Jamaica girls usually outperform boys on all school tasks at that age. He found the same results in a second Jamaican study. Better performance by boys on the same test did not occur in an American sample until Grade 9. Boe (1968) found girls in Grades 8, 10, and 12 to be less able to draw plane sections of solids than boys but equally able to choose the correct diagram of the section in a multiple choice test. Werdelin also found a sex-related difference in his Geometrical Construction test ($p < .10$) but it was not as significant as those in his two tests of problem solving in geometry.

In summary it seems that visual images may be used in mathematical problem solving to record the relationships among elements of the problem. Some problem solvers visually transform these elements into new combinations to arrive at a solution. Others rely more on verbal or

mathematical symbols to represent the information and transform these logically, sometimes very successfully. However, one has only to try to verbalize all the information contained in a geometric diagram to realize the economy of the visual representation.

As Werdelin (1961) pointed out, people who have both verbal and visual methods available are more likely to solve problems than those with only one method. The question of whether facility in one mode forecloses facility in the other will be left for other reviewers, but this review at least suggests that spatial training can improve problem solving performance. Given the evidence discussed in the first section of this chapter of the importance of knowledge of mathematical concepts and attitudes toward problem solving, spatial training should not be expected to eradicate sex-related differences in mathematical problem solving. However, it might remove an important component of those differences.

Technical Report No. 387
(Part 2 of 3 Parts)

THE INTERRELATIONSHIP OF SEX, VISUAL SPATIAL ABILITIES,
AND MATHEMATICAL PROBLEM SOLVING ABILITY IN GRADE SEVEN

by

Ann Koch Schonberger

Report from the Project on Conditions of School
Learning and Instructional Strategies

Thomas A. Romberg
Faculty Associate

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

June 1976

JUN 9 1980

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in the University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

WISCONSIN RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING

MISSION

The mission of the Wisconsin Research and Development Center for Cognitive Learning is to help learners develop as rapidly and effectively as possible their potential as human beings and as contributing members of society. The R&D Center is striving to fulfill this goal by

- conducting research to discover more about how children learn
- developing improved instructional strategies, processes and materials for school administrators, teachers, and children, and
- offering assistance to educators and citizens which will help transfer the outcomes of research and development into practice

PROGRAM

The activities of the Wisconsin R&D Center are organized around one unifying theme, Individually Guided Education.

FUNDING

The Wisconsin R&D Center is supported with funds from the National Institute of Education; the Bureau of Education for the Handicapped, U.S. Office of Education; and the University of Wisconsin.

ACKNOWLEDGEMENTS

Academic scholarship is not always the lonely pursuit it is often thought to be. Writing this acknowledgement provided me with an opportunity to recall the many pleasant relationships established during the research and writing of this dissertation.

First, I wish to thank my dissertation committee. Dr. Thomas Romberg, my major professor, first interested me in mathematical problem solving and assisted me greatly in the design, measurement, and data analysis aspects of my study. Dr. Elizabeth Fennema, whose interest in the effects of sex and spatial ability on mathematics achievement inspired my own interest, provided detailed criticism of this manuscript. Dr. Ann Becker contributed an important outside perspective to the study. I am grateful to all of them for their personal, as well as intellectual, support. Dr. Romberg allowed me to complete this study and other degree requirements in a rather unconventional way. Dr. Fennema and Dr. Becker provided examples of academic careers compatible with motherhood. Dr. Thomas Carpenter, Dr. John G. Harvey, and Dr. Mary Montgomery also read the manuscript and offered helpful comments.

Exchanging ideas about mathematical problem solving with other graduate students was useful to me in formulating my own ideas. The work of Ruth Ann Meyer, Diana Wearne, Don Whitaker, and Don Zalewski

was particularly important. In addition, Diana Wearne and Clement Steele took time from their own studies to code some of the problem solving tests.

The staff of the Bangor Public Schools cooperated by providing subjects for the study. Superintendent Wendell Eaton gave permission for the study. Mr. George Voss, Director of Research and Evaluation, made the necessary arrangements and provided some of the data. Mr. Lionel Lord and Mrs. Nancy Luttrell helped with the testing at the Fifth Street School. Thanks also go to the principals, the mathematics teachers, and the seventh-grade students at the Fifth Street and Union Street Junior High Schools who gave their time and cooperation.

I would also like to thank those who assisted me in the data analysis phase of my study. Dr. Jacob Evanson helped plan the analysis and Ms. Linda Junker assisted in the actual programming. The key-punching and computer operating staff of the Wisconsin Research and Development Center for Cognitive Learning provided competent and efficient service. Mrs. Dorothy Egner, also on the Center's staff, did the final typing.

Finally, I acknowledge the contributions of my family: my parents, my husband, and my children. My parents encouraged my intellectual activities and provided me with a first-class education. My children tolerated my absence in good humor. My husband unselfishly gave up his own research time to care for them as well as providing the emotional and financial support without which this project would have been impossible.

TABLE OF CONTENTS

	page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	vii
LIST OF FIGURES.	x
ABSTRACT	xii

Chapter

I	INTRODUCTION	1
	The Purposes of the Study	3
	Definitions	4
	Outline of the Thesis	7
II	REVIEW OF RESEARCH ON VISUAL SPATIAL ABILITIES .	9
	Introduction	9
	Spatial Subfactors.	10
	Sex-Related Differences in Visual Spatial	
	Abilities.	22
	Another Type of Spatial Test.	28
	An Alternate Framework.	34
	Causes of Sex-Related Differences in Spatial	
	Abilities.	41
III	REVIEW OF RESEARCH ON MATHEMATICAL PROBLEM	
	SOLVING ABILITY.	45
	Introduction.	45
	Sex-Related Differences in Mathematical	
	Problem Solving Ability.	46
	The Relationship Between Mathematical	
	Problem Solving Ability and Visual	
	Spatial Abilities.	65

Chapter		page
IV	DESIGN OF THE STUDY	89
	Introduction	89
	Rationale for This Study	89
	Selection and Description of the Sample.	92
	Selection of the Tests of Visual Spatial Abilities	95
	Construction of the Test of Mathematical Problem Solving	103
	Scoring Procedures	108
V	CARRYING OUT THE STUDY.	110
	Introduction	110
	The Pilot Study and Resulting Changes.	110
	Gathering the Data for the Main Study.	117
	Scoring Procedures	118
VI	DATA ANALYSIS	131
	Introduction	131
	Sex-Related Differences in Performance on the Tests	132
	Relationships Between Visual Spatial Abilities and Mathematical Problem Solving Ability and Sex-Related Differences in These Relationships	150
VII	SUMMARY AND CONCLUSIONS	183
	Introduction	183
	Summary.	183
	Conclusions of the Study	184
	Limitations of the Study	193
	Suggestions for Future Research.	196
	Educational Implications	197
	Conclusion	198
	REFERENCES	200
	APPENDIX A: Sample Items from a Number of Figural Tests . .	215
	APPENDIX B: Tests Used in This Study.	226

	page
APPENDIX C: Descriptive Statistics from Other Administrations of the Tests Used in This Study	248
APPENDIX D: Scatterplots of Two Methods of Scoring the SR-O Tests.	252
APPENDIX E: Correlation Matrices.	253

LIST OF TABLES

Table		page
2.1	<u>t</u> Ratios for Sex Differences in Spatial Test Means in the Project Talent Study.	25
3.1	Sex-Related Differences in Applications Scales, National Longitudinal Study of Mathematical Abilities, X-, Y- and Z-Populations.	56
3.2	Sex-Related Differences by Analysis Scales, National Longitudinal Study of Mathematical Abilities, X-, Y- and Z-Populations.	56
3.3	Correlations Among Three Abilities in the CAA Project (1970): Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	67
3.4	Correlations Among Three Abilities in the CAA Project (1971): Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	69
3.5	Correlation Coefficients of Spatial Tests and Analysis Scales in the NLSMA	71
3.6	Per Cents of Total NLSMA Analysis and Application Scales on Which Significant Main Effects of Spatial Variables Were Observed.	74
3.7	Relationships Among Three Abilities in the Project Talent Study: Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	77
3.8	Correlation Coefficients with Concealed Figures and Problem Solving in Students at Swarthmore and Yale	80
5.1	Descriptive Statistics from the Pilot Study Results.	112
5.2	Descriptive Statistics for Several Different Eighth-Grade Groups Taking the Spatial Tests	113
5.3	Textbook and Geometry Instruction for Each of the Mathematics Classes in the Sample.	119
5.4	Agreement over Coders 1, 2 and 3 on Solutions. . . .	125

Table		page
5.5	Agreement over Coders 1, 2, and 3 on Diagrams. . .	127
5.6	Agreement over Coders 1, 2, and 3 on Correct Representations.	129
6.1	Descriptive Statistics and Comparison of Means of Boys, Girls and Both for the Entire Test Battery	134
6.2	Reliability Estimates for Spatial Tests and Problem Scales for Boys, Girls and Both.	136
6.3	Problem Solving Item Analysis: Proportion of Correct Responses for Boys, Girls and Both	138
6.4	Distribution of Boys and Girls in Levels of Mathematics Classes.	140
6.5	<u>F</u> Values Generated by the Sex-by-Level ANOVAs. . .	142
6.6	Distribution of Boys and Girls in IQ Stanines. . .	145
6.7	<u>F</u> Values Generated by the Sex x IQ ANOVAs (Equal Cell Sizes).	147
6.8	Cell Means for the DAT Space Relations Test: Comparison of the Sex-by-IQ Subsample and the Whole Group.	148
6.9	Cell Means for the Cubes Comparison Test: Comparison of the Sex-by-IQ Subsample and the Whole Group.	149
6.10	The Relative Importance of Using Diagrams in Solving Three Types of Problems	154
6.11	Significant Sex-Related Differences in Correlation Coefficients	156
6.12	Correlations Between the Spatial Variables and Problem Solutions Total for the Whole Group. .	158
6.13	The Solutions Factor	161

Table		page
6.14	The Space Factor	163
6.15	The First Drawing Factor (Problem Types A and B) .	165
6.16	The Second Drawing Factor (Problem Type C)	167
6.17	Standardized Regression Coefficients	169
6.18	Standardized Regression Coefficients: Two- and Three-Dimensional Comparison	179
6.19	Canonical Correlation Analysis: Spatial Composite with Problem Solutions Composite	181

LIST OF FIGURES

Figure		page
1.1	Two Representations of a Rectangular Parallelepiped.	6
2.1	Guilford's Structure-of-Intellect Model	14
2.2	Figural Plane of the Structure-of-Intellect Model	14
2.3	A Translation in 2-space.	36
2.4	A Rotation in 2-space	36
2.5	A Reflection in 2-space	36
2.6	A Translation in 3-space	36
2.7	A Rotation in 3-space	36
2.8	A Reflection in 3-space	36
2.9	An Item Typical of the Cubes Test	39
4.1	A Matrix of Spatial Tests	96
4.2	A Set of Card Rotations Items	98
4.3	Two Cubes Comparison Items.	99
4.4	An Item from Form Board 2	100
4.5	An Item from DAT Space Relations.	102
4.6	A Matrix of Spatial Ability Tests Used in This Study	102
4.7	A Hidden Figures 2 Item	103
5.1	Example from the Instructions for the Cubes Comparison Test	115
5.2	An Example of a Figure Scored One for the Existence of a Diagram.	121

Figure		page
5.3	A Difficult Diagram to Score.	123
5.4	Some Solutions for Problem 6.	126
6.1	Sequence of Data Analysis, Part A, Sex-Related Differences in Performance.	133
6.2	Sex-by-Level Interaction: Cell Means on Diagram T ANOVA Using All Data.	143
6.3	Sequence of Data Analysis, Part B	152

THE INTERRELATIONSHIP OF SEX, VISUAL SPATIAL ABILITIES,
AND MATHEMATICAL PROBLEM SOLVING ABILITY IN GRADE SEVEN

Ann Koch Schonberger

Under the supervision of Professor Thomas A. Romberg

The intent of this study was to examine the hypothesis that males are more successful at solving mathematical and spatial problems than females. The relationship between visual spatial abilities and mathematical problem solving ability was also investigated. Specifically five questions were posed.

1. Are there sex-related differences in problem solving ability in mathematics?
2. Are there sex-related differences in any of the measures of visual spatial abilities?
3. What are the relationships between measures of visual spatial abilities and mathematical problem solving ability? Do these relationships differ by sex?
4. Are the relationships between visual spatial abilities and mathematical problem solving ability different for problems requiring differing amounts of spatial knowledge?
5. Is a subject's use of diagrams when solving mathematical problems related to successful solution of mathematical or spatial problems?

The sample consisted of the seventh grade students from one junior high school in Bangor, Maine. Each student took five spatial tests and a test of mathematical problems. The spatial tests were Card Rotations, Form Board 2, and Hidden Figures 2 from the NLSMA battery, Cubes Comparison from the ETS Kit, and Space Relations from the Differential Aptitude Test. The test of mathematical problems was constructed by the author from an item pool developed by Zalewski. Problems were chosen to make three subtests of problems differing in amount of spatial or geometric content. Level of mathematics class and IQ were also recorded for each subject. Questions about sex-related differences in performance were investigated using the t test; additional refinement was made possible by analysis of variance using the data on IQ and level of mathematics class. The relationships between solving spatial problems, drawing diagrams, and solving mathematical problems were explored using a variety of correlational techniques; these included factor analysis, multiple regression analysis and canonical correlation analysis.

Chapter IV

DESIGN OF THE STUDY

Introduction

A review of the research on visual spatial abilities and mathematical problem solving provides only partial answers to the questions raised in the introductory chapter. In this chapter the rationale for this study identifies and justifies the areas to be investigated. The selection of subjects, the selection and development of test instruments, and the plans for scoring the tests are also discussed in this chapter.

Rationale for This Study

Since studies cited in the previous chapters show that sex-related differences in visual spatial abilities including visual disembedding as well as in mathematics achievement begin to appear in Grades 6, 7, and 8, further study of this age group would be useful. One of the confounding variables in any study of problem-solving performance is the number and content of mathematics courses taken. Studying subjects in middle or junior high school where mathematics is required and tracking is less common would minimize this problem, as most students have had the same program.

Since these are the years when students undergo biological and perhaps intellectual changes, it is an important period to investigate the development of sex-related differences in visual spatial abilities and mathematical problem solving ability.

Middle-school or junior-high-school students were chosen for this study for another reason. In a large-scale study directed by Fennema and Sherman at the University of Wisconsin-Madison, visual spatial ability was one of the variables being studied as a possible correlate of sex differences in mathematics learning in Grades 9 through 12. In addition, Meyer's concurrent study of cognitive abilities (including spatial ability) and mathematical problem solving in fourth-grade students investigated sex differences. Upon completion of this study and those by Fennema and Sherman and by Meyer, a fairly comprehensive picture of the existence and development of sex-related differences in visual spatial abilities and mathematical problem solving ability should be available.

None of the studies mentioned in the preceding chapter examined correlations between problem-solving skill and a full range of measures of visual spatial abilities. This study was designed to do that. Both the SR-O, Vz division and the dimensional division discussed in Chapter II are investigated as well as visual disembedding or "field independence." Also consideration of different types of problems was either totally lacking in most

of the previous research, or else it was done only after the fact. Dodson's (1972) work is an exception but he studied eleventh-grade subjects and did not report on sex differences. The NLSMA is another exception but the spatial tests were given in a different year than the problem-solving tests and the relationships were not fully reported.

Specifically this is a status study using correlational techniques. The questions posed in the introductory chapter are investigated for a group of junior-high-school students. These questions can be restated as hypotheses and elaborated upon on the basis of the review of related research.

1. Boys and girls do not differ in their ability to solve mathematical problems.
2. There are no sex-related differences in performance on measures of any of the visual spatial abilities: SR-O or Vz, two- or three-dimensional, visual disembedding.
3. There are significant positive relationships between each of the visual spatial abilities and mathematical problem solving ability. The relationships are stronger for Vz tests than for SR-O tests and stronger for three-dimensional tests than for two-dimensional tests. These relationships do not differ by sex.
4. Each type of visual spatial ability is more closely related to solving mathematical problems with high spatial content than to solving problems with little spatial content.
5. Boys and girls do not differ in their use of diagrams in solving mathematical problems. Use of a diagram in solving a mathematical problem is positively related both to the ability to solve that problem and to visual spatial abilities. There are no sex-related differences in the relationships.

Selection and Description of the Sample

A population can be defined as the total group of elements or individuals about which one hopes to generalize. Populations can be real: the set of all seventh-grade students now enrolled in the Chicago public schools. They are often hypothetical: the set of all seventh-grade students. Much of the educational research dealing with hypothetical populations is concerned with differences between two groups such as seventh-grade boys and seventh-grade girls. Although the principal techniques of statistical inference are predicated on the assumption of random selection of the sample, for administrative reasons random sampling is often difficult. For hypothetical populations it is impossible, and yet statistical inferences are made about these populations. The rationale for violating the assumption of random selection is that the group studied (seventh-grade students in a certain school, for example) can be considered to be a random sample of a hypothetical population of seventh-grade students in schools like the one chosen. The problem of inference is then left to the reader. For this reason it is essential to specify in some detail the relevant characteristics of the group studied and the way in which the group was chosen. (Clark, Coladarci, & Caffrey, 1956)

The entire seventh-grade class of the Fifth Street Junior High School in Bangor, Maine, was selected as the sample for this study. As the second largest city in Maine, Bangor's population

is among the most heterogeneous in the state. Neighborhoods in Bangor are small enough so that each junior high school encompasses a number of different socio-economic levels. It is the author's opinion as a resident of Bangor that its population is fairly conservative with respect to sex roles. However a citizens' committee was appointed by the superintendent to study sex roles in the public schools. Reports on discrimination in course offerings (primarily home economics and manual arts) and in athletic programs were presented by the committee to the school board in 1974 and 1975, respectively. A few changes have resulted.

Of the three junior high schools in Bangor, Fifth Street was chosen as the most representative by the director of testing and research for the Bangor school system. He also provided additional information on the Fifth Street students as a whole to better describe this sample. In the Bangor school system the standardized test data is gathered on the eighth-grade students instead of the seventh, but the director's experience is that it varies little from year to year. The mean IQ of Fifth Street's eighth-grade students tested in 1974 was 104 measured by the Otis-Lennon Form J. The mean total mathematics score of the Metropolitan Advanced Battery Form F was 68 which is in the 42nd percentile. Mean scores of Fifth Street's student population on standardized IQ and achievement tests always fall between those of the other two junior high schools in Bangor, according to the director (Vose, G., Personal communication, June 1975).

Some estimate of the socio-economic and ethnic makeup of the Fifth Street School population can be gained from the following information. Fifth Street is not a Title I school. The only other socio-economic indicator available is the number of students who receive free or reduced-price lunches. At the junior-high level even this is not a good indicator since there is some social stigma attached; only 15% of the Fifth Street students participate. A better estimate (25%) can be obtained by combining the statistics for the two elementary schools which feed into this junior high school. The experimenter's observation during the testing sessions, which was supported by the school's guidance counselor, was that there were no Blacks, Chicanos, or Native Americans in the sample. Two subjects were Asian Americans. Maine's Franco-American minority which predominates in some areas of the state is not very substantial in Bangor.

The choice of seventh grade rather than eighth was made on the basis of the school's mathematics program. By the end of the eighth grade, some students have taken two semesters of algebra, some one semester, and some none at all. Since a problem for some one without algebra skills may be a routine exercise for a subject familiar with writing and solving equations, it seemed wiser to choose seventh-grade students instead. Although the seventh-grade mathematics program has two ability tracks, essentially the same material is the focus of instruction. Assignment to Level 1 (upper track) or Level 2 (lower

track) had been made on the basis of the Otis-Lennon score, recommendations of sixth-grade teachers, and score on the mathematics section of the Metropolitan Achievement Test also given in sixth grade. There were about 175 students enrolled in the seventh grade at Fifth Street in 1974-75.

Selection of the Tests of Visual Spatial Abilities

A number of different considerations entered into the choice of the tests of visual spatial abilities. The review of the literature indicated that factor analysis of spatial and mathematical tests often yields sex-related differences in factor structure. Investigation of factor structure was planned as one of the ways to investigate the questions raised in the introduction. Since factor analysis works best with large samples, the tests for the study had to be easily scored paper-and-pencil tests. A number of tests were to be given to each subject, and time was limited by the school system, so the tests had to be relatively short. Whenever advisable, tests used in other studies of visual spatial abilities and mathematical problem solving ability were used to facilitate comparison of results. In order to investigate both the SR-0, Vz division and the dimensional division, a two-by-two matrix was constructed and a test chosen for each cell. (See Figure 4.1.)

	SR-0	Vz
two-dimensional		
three-dimensional		

Figure 4.1. A matrix of spatial tests.

With one exception the tests of visual spatial abilities were chosen from the Kit of Reference Tests for Cognitive Factors (Revised, 1963) developed by the Educational Testing Service (referred to subsequently as the ETS Kit). The ETS Kit was developed for use in factorial studies of cognitive abilities and achievement. For each factor three tests were selected or adapted to cover as much of the range from sixth grade through college as possible, to be as factorially pure as possible, and to be as different as possible to balance out uniqueness. Easy administration by paper-and-pencil methods was another criterion. (French, Ekstrom, & Price, 1969)

Some of the tests from the ETS Kit were modified by the test development staff of the NLSMA to be used as part of their cognitive processes battery. Two general changes were made. Each ETS Kit test has two separately timed, parallel parts, whereas the corresponding NLSMA test uses just one of the parts. Forced choice tests from the ETS Kit are to be corrected for guessing and the instructions for each test reflect this. The corresponding NLSMA tests are scored

on number correct and the instructions only caution against guessing wildly. (Romberg & Wilson, 1969)

Choosing tests appropriate for seventh-grade subjects was somewhat difficult. Many spatial tests were originally designed to be used in vocational counseling, usually done with older students. Some of the tests from the ETS Kit were labeled appropriate for Grades 8 through 16. In the NLSMA most of the spatial tests were given in the fall of eighth grade. However, since the subjects in this study had nearly finished seventh grade, and the items required no reading, and the instructions would be read aloud, these tests were judged to be acceptable pending results of the pilot study.

The SR-O Tests

The ETS Kit manual calls this Factor S: Spatial Orientation. It is described as "the ability to perceive spatial patterns or to maintain orientation with respect to objects in space" and is "perhaps best thought of as space with the observer himself as a reference point" (French, Ekstrom, & Price, 1969, pp. 38-39). The figure is reacted to as a whole and may be translated, rotated, or reflected. These tests are usually speeded.

The two-dimensional test chosen was Card Rotations from the NLSMA battery which is very similar to Part 1 of the Card Rotations Test of the ETS Kit. (For a detailed description of the differences between the ETS and NLSMA tests, see Romberg and Wilson, 1969, page 184.)

The ETS Kit test, in turn, was based on Thurstone's Cards test and is like Thurstone's Flags and Figures. Each set of items shows a card cut into an irregular shape. Eight other drawings of the same card, either rotated or reflected, are shown. The task is to mark the drawings which show a rotation with a plus and those which show a reflection with a minus. Answers are recorded on the test itself. The entire test is included in Appendix B. For an example, see Figure 4.2.

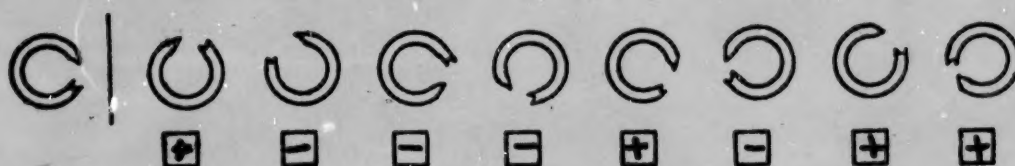


Figure 4.2. A set of Card Rotations items.

Card Rotations has 112 items and allows the subject four minutes. It was given in the NLSMA five times to students from fifth through eleventh grade. Descriptive statistics for these administrations are reported in Appendix C. Since this is a speed test and only one part was given to each group, no reliability coefficients were reported. Card Rotations was chosen for this study because it had been given to a large number of students of similar age, and its construct validity for this ability was high both logically and statistically.

In addition, it is closely related to Thurstone's tests used in studies reported in Chapter II.

The three-dimensional test chosen was the Cubes Comparison test from the ETS Kit which was adapted from Thurstone's Cubes. Each item presents two drawings of cubes with a letter, numeral, or other symbol on each face. Assuming no symbol can appear on two faces of the same cube, the subject is to indicate which pairs of drawings can represent the same cube and which pairs must represent different cubes. Answers are recorded on the test itself. For an example, see Figure 4.3. A copy of the test can be found in Appendix B. There are two pages, each with 21 items and each timed at three minutes, representing parallel forms. Statistics available from previous administrations of this test are reported in Appendix C.



Figure 4.3. Two Cubes Comparison items.

The Vz Tests

In the ETS Kit manual, this factor is called Factor Vz: Visualization. It is described as the "ability to manipulate or

transform the image of spatial patterns into other visual arrangements. Tests of this factor require the examinee to rotate, turn, fold, or invert the images of objects or parts of objects according to explicit directions and to make comparisons of the resulting manipulated images with drawings" (French, Ekstrom, & Price, 1969, p. 47). Although these tests are timed, speed is less of a factor than in SR-O tests.

The two-dimensional test chosen was Form Board 2 from the NLSMA battery which is very similar to Form Board Test, part 1, of the ETS Kit. (For differences, see Romberg and Wilson, 1969, page 185.) It is similar to Thurstone's Form Board and somewhat similar to the Revised Minnesota Paper Form Board Test by Likert and Quasha. Each item shows five shaded pieces, some or all of which can be put together like puzzle pieces to form the figure presented in outline form. The task is to check off which pieces must be used. According to the instructions, the pieces can be rotated and translated but not reflected. Answers are recorded on the test itself. For an example, see Figure 4.4.

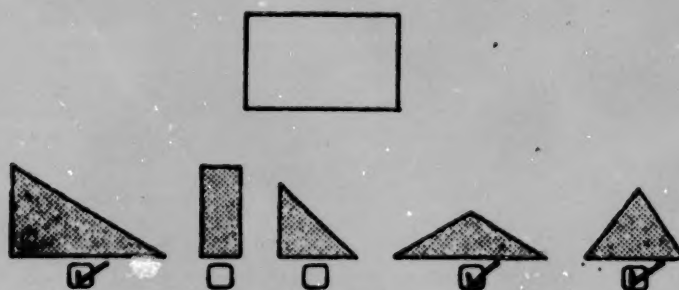


Figure 4.4. An item from Form Board 2.

Form Board 2, which can be found in Appendix B, has 24 items and allows the subject eight minutes to complete it. It was given twice by the NLSMA; coefficients of internal consistency of .75 and .76 were reported. (See Appendix C for descriptive statistics.) The choice of this test over the Revised Minnesota Paper Form Board Test, which is longer and may be more reliable, was made because of time. The ETS Kit and NLSMA battery contained another two-dimensional Vz test, Punched Holes. Form Board 2 was chosen for this study instead, because a test similar to it was being used by Meyer in her study of mathematical problem solving mentioned earlier.

The three-dimensional Vz test chosen was Space Relations from the Differential Aptitude Test battery (DAT) (Bennett, Seashore, & Wesman, 1973). Although it is longer than the other tests, the researcher wanted to use it in order to be able to compare her results with those obtained by Fennema and Sherman in the study mentioned earlier. The test seemed to be a logically valid measure for the Vz factor and the tasks are similar to those in the ETS Kit Vz test, Surface Development. The task in the DAT Space Relations test is to choose which of four drawings represents the figure which would be formed by folding and pasting the paper pattern shown. Answers are recorded on an IBM answer sheet. For an example, see Figure 4.5. This test has 60 items and takes 25 minutes. Form T is the only one available for separate administration. Reliabilities were reported between .93 and .95. Other descriptive statistics from the manual are reported in Appendix C.

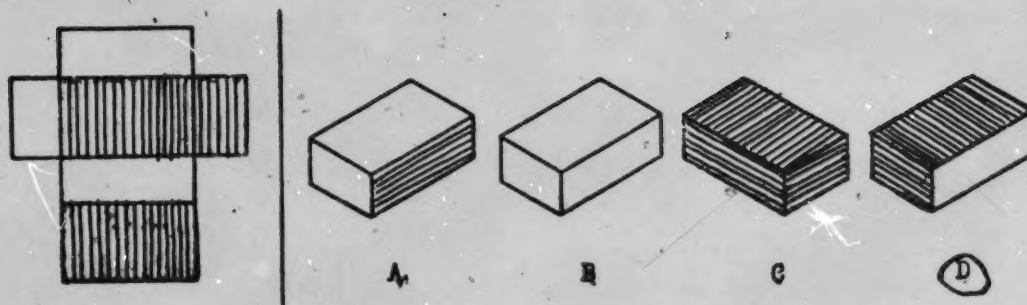


Figure 4.5. An item from DAT Space Relations.

Now entries can be made in the matrix described at the beginning of this section of the chapter. See Figure 4.6.

	SR-0	Vz
two-dimensional	Card Rotations	Form Board 2
three-dimensional	Cubes Comparison	DAT Space Relations

Figure 4.6. A Matrix of Spatial Ability tests used in this study.

The Test of Visual Disembedding

In the ETS Kit manual, this factor is called Factor Cf: Flexibility of Closure. It is defined as "the ability to keep one or more definite configurations in mind so as to make identification in spite of perceptual distractions" (French, Ekstrom, & Price, 1969, p. 9). According to the ETS Kit manual, it may be related to Witkin's construct of field independence and has some variance on the SR-0 factor. There are no three-dimensional tests of this factor in the ETS Kit.

Hidden Figures 2 from the NLSMA battery was chosen for use in this study. It is very similar to the Hidden Figures Test, part 1 of the ETS Kit. (For differences, see Romberg and Wilson, 1969, page 178.) The ETS Kit test is an adaption of the Gottschaldt Figures Test popularized by Thurstone. It is acknowledged to be a hard test. The task is to decide which of the five figures on the right is embedded in the more complex pattern at the left. Answers are recorded on the test itself. For an example, see Figure 4.7. This test has 16 items and takes 10 minutes. A copy is included in Appendix B. It was given in the NLSMA four times; descriptive statistics for these administrations are reported in Appendix C. Reliabilities ranged from .60 to .80.

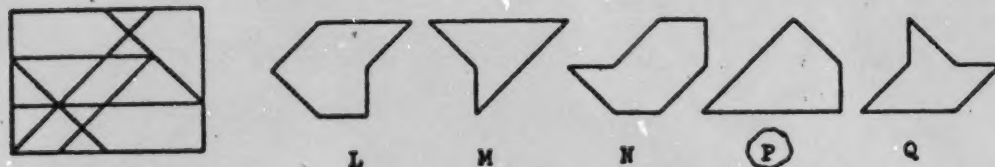


Figure 4.7. A Hidden Figures 2 item.

Construction of the Test of Mathematical Problem Solving

Choosing a valid and reliable written test of mathematical problem solving for seventh-grade students is itself a problem. The reading level and degree of computational difficulty of the problems must be appropriate. Only concepts understood by the subjects

should be used in the problems. Central to the validity of the test (and most difficult to assure) is that it is measuring real problem solving and not just routine applications.

Zalewski (1974) reviewed available procedures for assessing problem solving in seventh-grade students. None of the manuals of commercial tests surveyed offered a definition by which to judge its problems. Content validity was usually established by examining textbooks, study guides, and curriculum recommendations. The definition of a mathematical problem given in the introduction and used by Zalewski specifies that the individual attempting to attain the objective must not have a memorized answer or an immediate procedure. Problems for which seventh-grade students were not likely to have an immediate procedure were operationally defined by Zalewski as those requiring more than one step for solution, not containing direct verbal cues (such as sum, difference, total) for the operations necessary, and not providing a choice of answers. Few of the problems in the tests Zalewski surveyed satisfied these criteria.

Having demonstrated the need for a valid written test of mathematical problem solving, Zalewski constructed an item pool of 165 problems, any random sample of which could be used as a written test. Instead of basing his validity claims on the content of the written test, he attempted to establish the concurrent validity of the written test with an individually administered interview test. A panel judged

that the problems in the item pool for the interview test satisfied the definition of a mathematical problem. Each subject's performance on the interview test was assessed by means of a process-sequence coding system developed by Kilpatrick (1969) and refined by Lucas (1972). Zalewski assumed that what he was observing and measuring on the interview test was mathematical problem solving. Although he found that none of the correlations coefficients of scores on either of the two different written tests or their sum with the total score on the interview test reached his established minimum of .71, they were all above .60 and significant at the .001 level. He reported reliability coefficients above .80 for all the written tests and judged these to be satisfactory.

The decision to base the mathematical problem solving test for this study on problems from Zalewski's written test item bank was made for several reasons. In his study the written test performance did not quite account for 50% of the variance on the interview test but it came close, so the concurrent validity was considerable. As Zalewski cautioned, "the items in the written test could require one-step solutions and they need not meet the criteria for mathematical problems, but would attempt to avoid simple recall of knowledge or direct applications of algorithms" (Zalewski, 1974, pp. 50-51). In fact, content validity of the item pool appeared to be at least as high as that of commercial tests. Since this study was designed to include different types of problems as one of its dimensions,

selecting problems from a pool was preferable to using an intact test. Also Zalewski's item bank contains more geometric problems than most commercial tests.

In order to investigate the relationships between sex and visual spatial abilities and problems differing in amount of spatial content, three categories were established.

- A. Problems in which the stimulus (presentation of the problem) is partly pictorial or which require spatial or geometric skills or knowledge for solution.
- B. Problems with a completely verbal stimulus in which spatial skills (such as visualizing the situation or drawing a diagram) may be useful but are not necessary for solution.
- C. Problems which appear to have no spatial content. (In other words, any problems that do not fit in Categories A or B should be labeled C.)

Four judges categorized the entire written-test item pool. Two of the judges were teachers of college mathematics; one was male and one was female. The other two (one male and one female) were candidates for the Master's degree in mathematics and had had several years of experience teaching high school mathematics. A reduced item pool was formed of those problems assigned to the same category by all four judges.

In Category A a few of the items were quite similar to items from the tests of visual spatial abilities. Using these might have produced a spuriously high correlation between spatial ability

and mathematical problem solving ability. For this reason, these items (3, 4, 48, 143, 144) were eliminated from the pool.

Fundamental to solving a problem are understanding the concepts used and being able to do the necessary computation. Zalewski attempted to control the computational difficulty by limiting the computations to the four basic operations on non-negative rational numbers. Since this study was to be carried out at the end of the school year and most of the seventh-grade mathematics curriculum at the Fifth Street Junior High School was devoted to such computation, it was felt that this would not be a problem. This researcher identified the concepts and skills needed to solve the problems in the reduced item pool. Two of the five mathematics teachers at the school were interviewed to find out if their seventh-grade students would know these concepts and skills. A number of problems were eliminated, especially from Category A, since there had been very little work on geometry.

In order to sample as many different types of problems in each category as possible, the problems remaining after these three reductions were assigned to cells according to their algebraic structure or the task required. For example, three of the problems in Category C required comparing three or more ratios to find the best, largest, or fastest. In fact, most cells had only one problem. In order to construct a test of 24 problems, random selection of eight cells in each category was made. If there was

more than one problem in a chosen cell, a random choice was made within the cell.

Since sex-related differences were to be investigated in this study, it seemed advisable to control for sex bias in problem content (Federbush, 1974). Whenever possible, the problem was made neuter.

Examples: "Six girls belong to a basketball team. . . ." became

"Six students belong to a basketball team. . . ."

"A man owned three connected squares of land and wanted to divide it among his four children." became

"A mother and father owned three connected squares of land and wanted to divide it among their four children."

Where this was not possible, names and pronouns were changed so that there were equal numbers of male-centered and female-centered problems in each category.

The test was constructed with each problem at the top of one page followed by the direction "(You may draw any pictures or diagrams that you think may help you.)" A blank for the answer was placed at the bottom of the page. The order of the problems was randomly assigned but the same for all subjects.

Scoring Procedures

The plan was to score the tests of visual spatial abilities as they had been in previous use. For all but the Cubes Comparison test, the score was the number correct. In the Cubes Comparison test, the

score was the number right minus the number wrong. This plan was altered in the main study. (See Chapter V.)

Four scales were planned for the test of mathematical problem solving, each to be applied to each of the three subtests and the total test. The problem-solutions scale tells the number of correct answers. The diagrams scale counts the number of problems for which diagrams were drawn. The correct-representation scale indicates how many of the diagrams represent the information in the problem correctly. The percent scale is a comparison of the last two scales.

Chapter V

CARRYING OUT THE STUDY

Introduction

The plan for this study as described in the last chapter had already been influenced by some practical considerations particularly in the area of selecting the subjects. The pilot study and the changes it suggested are discussed in this chapter as is the data gathered in the main study. The detailed procedures for scoring the tests are also discussed.

The Pilot Study and Resulting Changes

The pilot study had two main purposes. One was to see if the tests were appropriate for seventh-grade subjects, since statistics had been reported only for eighth-grade students and the means reported were low. The other purpose was to plan the time schedule for administering the tests in the main study. The plan was to give the spatial tests in two sessions and the test of mathematical problem solving in one session. The principal at the Union Street Junior High School in Bangor agreed to provide three seventh-grade mathematics classes for the pilot study. Since the mathematics classes at Union Street are heterogeneously grouped, it was assumed

that a full range of ability would be available in any one class. As it turned out, one class participated in two sessions and the other in one. The experimenter administered all the tests, reading aloud the printed instructions accompanying each test while the subjects read them silently. The tests were scored as described in the preceding chapter.

All the tests of visual spatial abilities are timed tests so the main concern was the time necessary for giving instructions. The tests selected for the first session were the ones from the NLSMA battery. They were given in this order with the time allowed for solving items in parentheses: Card Rotations (4 minutes); Form Board 2 (8 minutes); Hidden Figures 2 (10 minutes). Since Card Rotations is an easy test, it was given first. The order of the other two tests was determined by the transformations allowed. In Card Rotations the task calls for translating, rotating, or reflecting figures. In Form Board 2 translating and rotating are allowed but not reflecting. In Hidden Figures 2 only translating is permitted. Cubes Comparison (6 minutes) and DAT Space Relations (25 minutes) were given in that order at the second session. Both sessions were completed easily within the 50-minute period. An hour was allowed for the test of mathematical problem solving. All students finished within 50 minutes and the average time to completion was 39 minutes.

Judging whether or not the tests were appropriate to seventh-grade students was more difficult. Table 5.1 presents the pilot study's results. Reliability coefficients for Form Board 2, Hidden Figures 2, and DAT Space Relations were estimated using Cronbach's α which is equivalent to the Kuder-Richardson Formula 20 when the scoring is one-zero. The reliability of the Cubes Comparison test was estimated by computing a Pearson r between the two separately timed pages and correcting it using the Spearman-Brown formula. Since Card Rotations is a speeded test, estimating reliability by any coefficient of internal consistency is inappropriate. There was not enough time to give parallel forms of this test, so the fact that many other researchers have used the test was accepted as a measure of its reliability. Since none of the tests were finished by all subjects, the reliability coefficients should be considered to be overestimates of the real reliability of the tests for these subjects.

Table 5.1
Descriptive Statistics from the Pilot Study Results

Test	Number of Subjects	Mean	Range	SD	Reliability
Card Rotations	24	50.48	68	18.36	
Form Board 2	24	5.17	13	2.97	.68 ^a
Hidden Figures 2	24	3.13	7	2.03	.58 ^a
Cubes Comparison	16	7.63	31	5.27	.47 ^b
DAT Space Relations	17	24.20	41	9.55	.89 ^a

^aCronbach's α

^bPearson r between parallel forms corrected by Spearman-Brown formula

With the exception of the DAT Space Relations tests, none of these reliability coefficients is impressive. On the other hand, all but DAT Space Relations are short tests, the samples were small, and the coefficients are similar to those reported for eighth-grade groups taking these tests. (Table 5.2 presents descriptive statistics for a number of different eighth-grade groups who have taken these tests. References for these and statistics for other age groups taking the tests can be found in Appendix C.) Since the tests were going to be used to investigate factor structure and to compare groups rather than to make educational decisions about individuals, the low reliabilities were judged acceptable. It was also predicted that the reliability estimates would be higher in the main study which would have more subjects.

Table 5.2

Descriptive Statistics for Several Different Eighth-Grade Groups
Taking the Spatial Tests

Test	Sample Size	Mean	SD	Reliability ^a
Card Rotations	1565	54.98	20.85	none reported
Form Board 2	565	6.65	3.65	.76
Hidden Figures 2	1091	4.37	2.58	.60
	1589	4.01	2.48	.59
Cubes Comparison	156	9.81	2.96	none reported
DAT Space Relations				
Boys	7000+	24.1	9.9	.94
Girls	6900+	23.0	9.0	.95

^aCronbach's α

Both the subjects' spontaneous remarks and the low means indicated that the tests were hard. On the other hand, the variability was generally high and both the means and standard deviations for each test were similar to those of eighth-grade groups. The only other measures available were the NLSMA tests Form Board 1 and Hidden Figures 1, which were designed for fourth- or fifth-grade students and might have been too easy. Since no other classroom-sized group was available to pilot test another set of measures, and these seemed the best available, the decision was made to use these tests and explore other ways of improving the performance of the subjects in the main study.

Informal observation of the subjects in the pilot study led the experimenter to suspect that a few subjects performed poorly because they did not understand the tasks, especially on the Cubes Comparison test. Changing the instructions would limit comparison with other groups. However, comparisons between this study and other groups had always been considered secondary to the comparisons within the group in this study. Since the changes under consideration were not substantial, it seemed that cautious comparison with other groups might still be valid.

When the NLSMA test development staff adapted Card Rotations, Form Board 2, and Hidden Figures 2 from the ETS Kit, they omitted the answer to the last example in each set of instructions (Romberg & Wilson, 1969). This experimenter decided to go over the last example on each test after the students had attempted it, providing answers

but being careful not to provide additional hints or strategies. During the pilot study some subjects apparently had difficulty understanding the difference between "turned around" and "turned over." Two large cards, a right triangle and a P-shape like those shown in the instructions, were prepared to be used to demonstrate the difference between rotation and reflection in the main study. The instructions to be read remained the same. The only other change for these three tests was in the instructions for Form Board 2. The experimenter planned to remark that these problems were like the wooden puzzles the subjects used to do and, just like those, these pieces could be turned around but not turned over.

In the pilot test of Cubes Comparison, some subjects appeared not to understand the real situation represented by the pictures of cubes. To overcome this, the experimenter constructed three cardboard cubes to be used in the main study. Two were marked like the left hand pair in the instructions so that it could be shown that there was no way to turn one so that it would look like the other. (See Figure 5.1.) The third cube was marked to show that



Figure 5.1. Example from the instructions for the Cubes Comparison Test.

the two pictures on the right in Figure 5.1 were different views of the same cube. No changes were made in the instructions for DAT Space Relations.

Since the test of mathematical problem solving had been newly constructed for this study, it was examined item by item after the pilot test. The subjects reported that the test was hard but that they could read and understand the problems. The mean, which was 7.11, was low but not unexpectedly so considering previous results with problems from this item pool (Zalewski, 1974). The variability was good: the range was 16 and the standard deviation was 5.21. Reliability estimated by the Kuder-Richardson Formula 20 was .89.

There were two items on the problem-solving test that caused trouble. On one the author disagreed with Zalewski's answer, and the subjects' responses split about equally between the two. There were other problems in the cell from which that one had been drawn, so another problem was randomly selected from that cell to replace it. In another problem several subjects got the right answer for the wrong reason, as their paperwork indicated. This problem was replaced by another from the same category.

Several changes were made in the cover sheet for the problem-solving test. During the pilot test, some subjects seemed not to be working on a problem very long before giving up. The instructions were changed slightly in hopes of promoting perserverance without increasing frustration. The cover sheet for the problem-solving

test originally had a blank for sex after the one for name of subject. There were a lot of jokes about what to write, and several students asked if the purpose of the study was to compare boys and girls. Since there is some evidence that females may try to avoid success in competition with males, especially in male-dominated fields (Horner, 1974), it seemed advisable to gather this information another way.

Gathering the Data for the Main Study

Before the testing began, the school's guidance director spoke to each seventh-grade mathematics class about the study and the schedule for testing. He was instructed not to mention that sex differences were being investigated when he told the students about the purpose of the study. The testing was done between May 27 and June 5, 1975 in the school's auditorium--study hall. Two mathematics classes from the same track or level were tested at once. All groups received the tests in the same order.

Session 1: Card Rotations; Form Board 2; Hidden Figures 2

Session 2: Cubes Comparison; DAT Space Relations

Session 3: Test of Mathematical Problem Solving

A very few students were absent for one or two test sessions. Most of these were given a makeup test later during this ten-day period. In the end there were very few missing scores: of 176 subjects four missed session one, none missed session two, and three missed session three. Identifying numbers were used to insure the students' anonymities.

Other data were collected for each student. The students wrote the name of their current mathematics teacher and their level on their test papers. Working from the master list of names and identifying numbers, the guidance office supplied sex of student and IQ in stanines measured by the Otis-Lennon Form J given in the fall of sixth grade. The seventh-grade mathematics teachers were interviewed to provide additional information about the subjects' mathematics programs in the year they were tested. This was done in order to have a rough estimate of the concepts and skills that students had had an opportunity to learn. All the teachers reported having spent most of the year teaching computation of non-negative rational numbers and its applications to the verbal problems. They were also asked what topics in geometry had been taught in that year. Table 5.3 presents a summary of this information and the textbook used in each class. It should be noted that three of the four Level 2 classes had no geometry in the 1974-75 school year. Teachers 1 and 4 are women; teachers 2, 3, and 5 are men.

Scoring Procedures

The plan described in Chapter IV was to score the tests of visual spatial abilities as they had been scored in previous use. Upon surveying the data from the main study, the experimenter noticed that there seemed to be little relationship between the scores on Card Rotations and Cubes Comparison which are supposed to measure

Table 5.3

Textbook and Geometry Instruction for Each of the Mathematics
Classes in the Sample

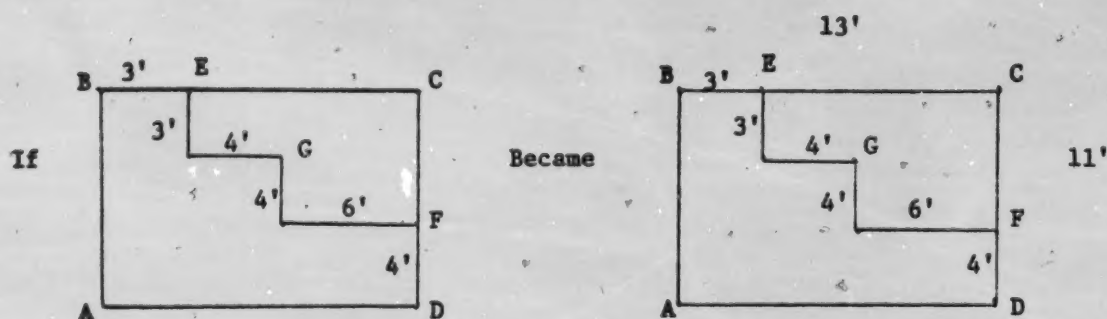
Teacher	Level	Textbook	Amount of Geometry Taught This Year
1	2	<u>Modern School Mathematics: Structure and Use</u> (Houghton-Mifflin)	Teacher-made worksheets on area, perimeter, and volume. Worksheets on geoboard. One week total.
2	2	<u>Modern School Mathematics: Structure and Use</u>	None
3	2	<u>Growth in Mathematics</u> (Laidlaw)	None
5	2	<u>Modern School Mathematics: Structure and Method</u> (Houghton-Mifflin)	None
2	1	<u>Modern School Mathematics: Structure and Method</u>	Definitions and basic relationships of lines, points, planes, angles.
3	1	<u>Mathematics: Structure and Use</u> (Science Research Associates)	Definitions, a little on area and perimeter.
4	1	<u>Modern School Mathematics: Structure and Use</u>	One week of transformational geometry. Less than a week on definitions. One week on area and perimeter of triangles, rectangles and parallelograms.
5	1	<u>Modern School Mathematics: Structure and Method</u>	Perimeter, area, and volume. One week.

the same factor. This was not noticed in the pilot study since two different classes were used in trying out the two groups of spatial tests and no correlation coefficients were computed. It seemed that the lack of relationship might be due to the fact that, although both tests are forced choice with two alternatives, one had been corrected for guessing and one had not.

Scoring both in the same way seemed advisable; the question was which way. The decision to correct both tests for guessing by subtracting the number wrong from the number right was made for several reasons. The instructions for Cubes Comparison had specifically stated "Your score on this test will be the number marked correctly minus the number marked incorrectly." The instructions for Card Rotations had only cautioned the subjects against guessing wildly. No subject had asked if Card Rotations would be corrected for guessing. It seemed either scoring procedure for Card Rotations was justifiable in terms of the instructions.

Another factor supported this decision. Scatterplots were prepared for each test comparing the corrected and uncorrected scores. (See Appendix D.) The scatterplot for Card Rotations was nearly a straight line; most subjects had very few items incorrect. The graph for Cubes Comparison was more scattered. Therefore, it was assumed that less change would occur in the scale statistics for Card Rotations by correcting for guessing. However, changing the scoring for Card Rotations further limits comparison with the NLSMA population data.

Applying the scoring system for the problem-solving test described in Chapter III to the pilot data raised several questions. The second scale was originally intended to measure the number of diagrams drawn by the subjects. However, in five problems a diagram is presented as part of the stimulus. Since use of a diagram was what was under investigation, it was decided to assign a score of one to those problems if the subject had further marked or modified the existing diagram in a rational way. For an example, see Figure 5.2. It was recognized that "in a rational way" necessitates judgments on the part of the scorer, but examination of subjects' responses indicated that some sifting and winnowing were advisable. Diagrams which had been drawn and then erased or crossed out were counted on the assumption that they had been used in some way in trying to solve the problem.



Problem: What is the perimeter of (distance around) the rectangle ABCD?

Figure 5.2 An example of a figure scored one for the existence of a diagram.

There were also some decisions to be made as to whether certain representations were in fact pictorial. The value of having defined all three types of representations in the introduction became apparent. If a given representation was judged not to be enactive and not to be symbolic, then it must be pictorial. Representations using numbers are usually symbolic, but there are two problems in which this was judged not to be the case. Problem 5 involves finding the number of linear orders of four objects (students in a lunch line). Many subjects used numbers in a nominal sense to identify the elements in the line; some used letters or other symbols such as circles, squares, or stars. Since the order and the line are the crucial elements of the representation, it was judged to be pictorial. A similar judgment was made regarding the use of numbers in the nominal sense in Problem 19.

Six students belong to the basketball team but only five can play at a time. How many groups of five players can be formed by six students?

Representations of this problem were labeled pictorial in which six symbols were arranged in groups of five or listed and crossed off apparently one at a time.

Decisions as to whether or not existing diagrams were correct were also subjective. Incomplete representations were judged correct if they were correct as far as they were drawn and the answer could logically follow from the diagram. One decision that gave the author some trouble was the one pictured in Figure 5.3.

Of the 13 students who drew this diagram, seven wrote the correct answer. This representation was counted as correct with some misgivings. If more than one diagram was drawn and one was correct, a score of one was assigned.

Problem 1

There was half of an apple pie left. Then Nate ate one half of the half and Kate ate one half of what Nate left. What part of the pie was left after Kate ate?

Diagram



Figure 5.3. A difficult diagram to score.

All the scoring decisions discussed in the preceding paragraphs were made by the experimenter. These decisions substantially reduced the number of judgments to be made when scoring individual tests. Since there was still some subjectivity involved, agreement with other raters was desirable. Five tests containing problems difficult for the experimenter to score plus a random sample of 20 more were selected from the main study. Coder 2 and Coder 3, PhD students in mathematics education who were familiar with problem-solving research, independently scored these 25 tests. They were given a list of acceptable solutions, the definition of a diagram and a summary of the decisions discussed in the preceding paragraphs.

They practiced on two different tests with the experimenter before scoring the chosen 25 on their own. The author was Coder 1.

Inter-coder reliability was computed separately for the three scales: solutions, diagrams, and correct representations. Pairwise comparisons were made and ratios of agreement were computed for each problem. For correct representations the number varied according to how many diagrams had been used. The results of the pairwise comparisons were averaged to make up the tables presented below. (It should be noted that since the scoring is one-zero, absence of total agreement still means that two of the three coders agreed.) Reliability for each scale was estimated by averaging the agreement ratio averages over all 24 problems.

On the solution scale, there was total agreement across all three raters on 17 of the 24 problems. (See Table 5.4.) On six more problems the coders disagreed on only 1 of the 25 solutions; all of these could be attributed to coder error, perhaps due to fatigue or inattention. The remaining problem, Number 6, was difficult to score. The task was to divide an L-shaped piece of land made of three squares equally among four children. Any solution in which each child got three-fourths of the square was considered correct; the shares did not have to be the same shape nor did any one child's share have to be continuous. (See solutions B and C of Figure 5.4.) With solutions like D, E, or F unless the subject drew the dotted lines, it was very difficult to separate

Table 5.4

Agreement over Coders 1, 2, and 3 on Solutions

Problem	Average No. of Agreements	Average No. of Disagreements	Average of Agreement Ratios
1	24.3	0.7	.97
2	24.3	0.7	.97
3	25.0	0.0	1.00
4	25.0	0.0	1.00
5	25.0	0.0	1.00
6	21.3	3.7	.85
7	25.0	0.0	1.00
8	24.3	0.0	.97
9	25.0	0.0	1.00
10	25.0	0.0	1.00
11	25.0	0.0	1.00
12	25.0	0.0	1.00
13	25.0	0.0	1.00
14	24.3	0.7	.97
15	25.0	0.0	1.00
16	25.0	0.0	1.00
17	25.0	0.0	1.00
18	24.3	0.7	.97
19	25.0	0.0	1.00
20	25.0	0.0	1.00
21	24.3	0.7	.97
22	25.0	0.0	1.00
23	25.0	0.0	1.00
24	25.0	0.0	1.00
Average of Agreement Ratios over 24 Problems			.99

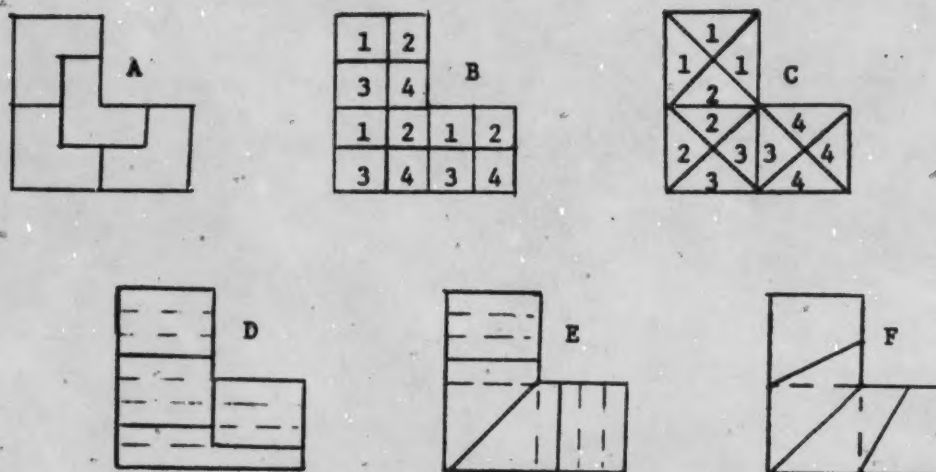


Figure 5.4. Some solutions for Problem 6.

the solutions from the rough guesses. There was considerable coder bias: Coder 3 was more lenient than Coder 2 who was, in turn, more lenient than Coder 1. Since there was disagreement on 6 of the 25 solutions and the experimenter had not recognized E or F as solutions, she decided to rescore that problem on all of the tests.

On the diagram scale there was total agreement on 12 of the 24 problems. (See Table 5.5.) On 8 more problems the Coders disagreed on only 1 of the 25 responses. All of these were attributed to Coder errors of different types; for example, three of the eight were cases in which a Coder did not count a crossed out or erased diagram as instructed. Problems 6 and 19 had a higher number of

Table 5.5

Agreement over Coders 1, 2, and 3 on Diagrams

Problem	Average No. of Agreements	Average No. of Disagreements	Average of Agreement Ratios
1	25.0	0.0	1.00
2	24.3	0.7	.97
3	25.0	0.0	1.00
4	24.3	0.7	.97
5	22.3	3.7	.89
6	23.0	2.0	.92
7	25.0	0.0	1.00
8	24.3	0.7	.97
9	25.0	0.0	1.00
10	25.0	0.0	1.00
11	25.0	0.0	1.00
12	24.3	0.7	.97
13	24.3	0.7	.97
14	24.3	0.7	.97
15	24.3	0.7	.97
16	25.0	0.0	1.00
17	25.0	0.0	1.00
18	25.0	0.0	1.00
19	23.7	1.3	.95
20	24.3	0.7	.97
21	25.0	0.0	1.00
22	25.0	0.0	1.00
23	25.0	0.0	1.00
24	21.7	3.3	.87

Average of Agreement Ratios over 24 Problems .97

disagreements but they still all appeared to be Coder errors. On Problem 5 three of the four disagreements involved lines drawn to represent the lunch line. Coder 1 scored them all one, Coder 2 scored them all zero, and Coder 3 scored two of them one for being purposeful and one of them zero which appeared to her to be random marks. Problem 24 caused difficulty because of the author's decision to score as one diagrams presented with the problems if they were marked or modified "in a rational way." Dots on the faces of the cubes seemed rational to Coder 2 more often than to Coder 1 and to Coder 1 more often than to Coder 3.

Agreement ratios on the correct representation scale were lower than those on the other two scales. (See Table 5.6.) Although a number of specific decisions had been made regarding the representations, it was recognized that scoring would be more subjective on this scale than on the others. On Problem 1 Coder 3 was more lenient in scoring the cutting of the pies than Coder 1, and Coder 1 was more lenient than Coder 2. On Problems 2 and 24 Coder 1 was more lenient but Coders 2 and 3 were equally severe. In Problem 13 Coder 3 was more particular about the representation of the tile shaped like an equilateral triangle than the other two raters. The issue in Problem 9 was labeling of the diagonal, $\overline{DE}$: Coder 3 allowed either 10 or 11, since it was not needed for solving the problem; Coder 1 allowed 11 because it was close to the correct answer of $\sqrt{125}$, but not 10 because $\overline{DE}$ seemed obviously longer than $\overline{EF}$ which was 10, Coder 2 allowed neither labeling. Since a correct

Table 5.6

Agreement over Coders 1, 2 and 3 on Correct Representations

Pro-blem	No. of Diagrams Observed	Average No. of Agreements	Average No. of Disagreements	Average of Agreement Ratios
1	19	17.00	2.00	.89
2	21	18.00	2.00	.86
3	0			
4	8	8.00	0.00	1.00
5	22	21.33	.67	.97
6	24	20.00	4.00	.83
7	10	10.00	0.00	1.00
8	0			
9	5	4.00	1.00	.80
10	0			
11	0			
12	1	.33	.67	.33
13	9	7.00	2.00	.78
14	1	.33	.67	.33
15	3	3.00	.00	1.00
16	0			
17	0			
18	0			
19	8	4.67	3.33	.58
20	0			
21	0			
22	0			
23	6	4.67	1.33	.78
24	9	9.00	.00	1.00
Average of Agreement Ratios over 24 Problems				.80

solution to Problem 6 was a correct representation and nothing more, the difficulty in scoring the solutions to Problem 6 was reproduced on this scale. The remaining problem with a low agreement ratio was Number 19. Several of the disagreements were due to different judgments on incomplete representations.

In general the inter-coder reliability seemed satisfactory to the author. With the rescoring of solutions to Problem 6, agreement on that scale should be very high. Similarly the ratings on the diagram scale seemed reliable. The reliability on the correct representation scale was lower so results involving that scale should be interpreted more cautiously.

Chapter VI

DATA ANALYSIS

Introduction

In the last chapter the administration and scoring of the tests were described as well as the gathering of other data on the students in the sample. The procedures for analyzing the data and the results of the analyses are given in this chapter. Although the sample was not randomly selected, it was considered to be a random sample of a hypothetical population. The justification for using sample statistics based on the assumption of random sampling is discussed in Chapter IV in the section on the selection and description of the sample. The limitations of this decision are discussed in Chapter VII.

In the introductory chapter a number of questions were posed: questions about differences in performance, questions about relationships, and questions about differences in relationships. These questions were restated as hypotheses in Chapter IV after the review of related research. Answering these three different types of questions required a number of different statistical methods. Also, some questions were investigated using more

than one statistical technique. The computer programs used in the data analysis were standard programs available from the Madison Academic Computing Center. STAT JOB programs were used for all but the canonical correlations; it was necessary to employ a FINN program for that analysis.

Part A. Sex-Related Differences in Performance on the Tests

This section deals with the investigations concerning sex-related differences in performance. Specifically, the first, the second, and part of the fifth hypotheses stated in Chapter IV are examined.

1. Boys and girls do not differ in their ability to solve mathematical problems.
2. There are no sex-related differences in performance on measures of any of the visual spatial abilities: SR-0 or Vz, two- or three-dimensional, visual disembedding.
5. Boys and girls do not differ in their use of diagrams in solving mathematical problems.

Additional refinement was possible using the data available on the IQ and level of mathematics class of each student. Figure 6.1 indicates the sequence of the data analysis discussed in this section.

1. Descriptive statistics. For all the tests of visual spatial abilities and the subscales derived from the test of mathematical problem solving, descriptive statistics were computed. These statistics are reported in Table 6.1.

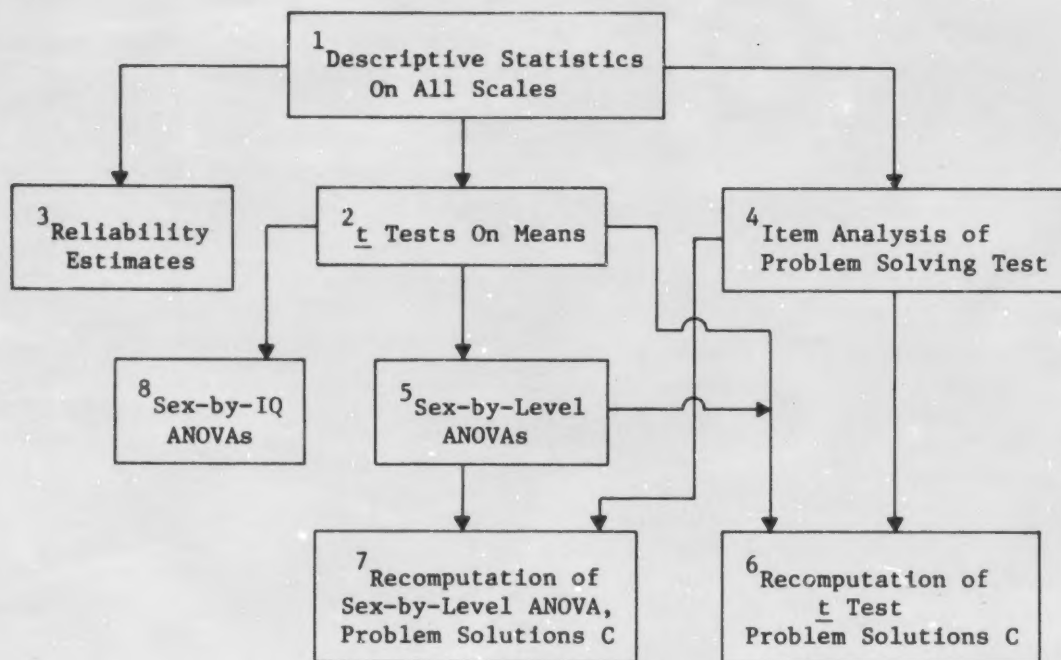


Figure 6.1. Sequence of data analysis, part A, sex-related differences in performance.

2. t tests on means. The significance of differences between the girls' and boys' means were determined using t tests and are reported in Table 6.1. Of all five of the spatial tests, only on Form Board 2 was there a significant sex-related difference in the scores; the difference was significant at the .01 level. Using the ω^2 statistic (Hays, 1973), sex accounted for about 5% of the variance on that test. Sex-related differences in favor of girls appeared on seven of the

Table 6.1

Descriptive Statistics and Comparison of Means
of Boys, Girls and Both for the Entire Test Battery

Tests	Number of Items	Ranges			Means			Standard Deviations			Mean Dif- ference t value
		Boys	Girls	Both	Boys	Girls	Both	Boys	Girls	Both	
1. Form Board 2 ^a	24	17	14	17	6.36	4.9	5.66	3.68	2.92	3.40	3.31**
2. Hidden Figures 2 ^a	16	9	10	10	2.42	2.72	2.56	1.76	2.38	2.08	-.93
3. DAT Space Relations ^b	60	41	45	52	26.67	26.07	26.39	9.69	9.48	9.57	.41
4. Problem Solutions A	8	8	7	9	1.67	1.67	1.67	1.45	1.41	1.42	.00
5. Problem Solutions B	8	8	9	9	2.09	2.19	2.14	1.76	1.93	1.84	.35
6. Problem Solutions C	8	9	9	9	3.34	3.00	3.18	1.93	2.02	1.98	1.13
7. Problem Solutions T	24	19	20	20	7.10	6.87	6.99	4.30	4.41	4.35	.35
8. Diagrams A	8	8	9	9	2.78	3.51	3.13	1.49	1.93	1.75	-2.76**
9. Diagrams B	8	5	5	5	1.78	2.16	1.96	1.03	.94	1.00	-2.53*
10. Diagrams C	8	2	2	2	.04	.12	.08	.21	.32	.27	-1.92
11. Diagrams T	24	11	13	13	4.60	5.78	5.17	2.17	2.56	2.43	-3.25**

Number of girls taking all tests = 83

^a Number of boys taking these tests = 89

^b Number of boys taking these tests = 93

Number of boys taking the remaining tests = 90

*p < .05

**p < .01

Table 6.1 (Continued)

Descriptive Statistics and Comparison of Means
of Boys, Girls and Both for the Entire Test Battery

Tests	Number of Items	Ranges			Means			Standard Deviations			Mean Dif- ference t value
		Boys	Girls	Both	Boys	Girls	Both	Boys	Girls	Both	
12. Correct Representations A	8	5	7	7	.89	1.22	1.04	1.03	1.41	1.23	-1.74
13. Correct Representations B	8	4	5	5	.78	1.11	.94	.78	.83	.81	-2.68**
14. Correct Representations C	8	2	2	2	.03	.08	.06	.18	.28	.23	-1.38
15. Correct Representations T	24	6	9	9	1.68	2.37	2.01	1.45	1.98	1.76	-2.59**
16. Per Cent A	100	101	101	100	28.23	31.26	29.70	29.67	30.32	29.94	-.66
17. Per Cent B	100	101	101	100	44.06	51.80	47.88	37.73	33.85	35.96	-1.42
18. Per Cent C	100	101	101	100	60.00	77.78	71.43	54.77	40.10	46.88	-2.35*
19. Per Cent T	100	101	101	100	34.27	38.77	36.44	25.98	25.87	25.95	-1.14
20. Card Rotations ^a	112	104	104	109	45.57	46.98	46.25	21.44	19.88	20.66	-.44
21. Cubes Comparisons ^b	42	40	50	50	8.67	6.83	7.80	8.69	8.28	8.52	1.43

Number of girls taking all tests = 83

^a Number of boys taking these tests = 89

^b Number of boys taking these tests = 93

Number of boys taking the remaining tests = 90

*p < .05

**p < .01

subscales of the test of mathematical problem solving related to use of a diagram: Diagrams A, B, and T, Correct Representations B and T, and Per Cent C. Again using the ω^2 statistic, sex accounted for 3% to 6% of the variance on each of these scales. It is interesting to note that there were no significant sex-related differences in the means of the Problems Solutions scales. This is contrary to results of other studies of seventh- and eighth-grade students reviewed in Chapter III.

3. Reliability estimates. The importance assigned to these results must be based on the reliability of the tests. (See Table 6.2.) Reliability was estimated using the Kuder-

Table 6.2

Reliability Estimates for Spatial Tests and
Problem Scales for Boys, Girls and Both

Tests	Boys	Girls	Both
1. Form Board 2	.78	.69	.75
2. Hidden Figures 2	.44	.69	.59
3. DAT Space Relations	.88	.88	.88
4. Cubes Comparison	.69	.69	.69
5. Problem Solutions A	.48	.42	.45
6. Problem Solutions B	.60	.69	.64
7. Problem Solutions C	.62	.69	.65
8. Problem Solutions T	.80	.81	.80

Richardson Formula 20 on Form Board 2, Hidden Figures 2, DAT Space Relations, and the four problem-solutions scales of the test of mathematical problem solving. Since Form Board 2 and Hidden Figures 2 are somewhat speeded, the reported coefficients of internal consistency were probably overestimates of the reliability of those two tests. A Pearson product-moment correlation coefficient was calculated using the two separately timed equivalent subtests of the Cubes Comparison test. Since the Pearson r is really an estimate of the reliability of the 21-item subtests, the Spearman-Brown formula was used to correct the estimate for the 42-item test. Since the content of each item on the test of mathematical problem solving determined to a considerable extent whether or not the subject would draw a diagram, coefficients of internal consistency were considered inappropriate for all the problem-test subscales concerned with diagrams.

4. Item analysis of the problem solving test. The test of mathematical problem solving was newly constructed for this study so an item analysis was performed on the results of this test. (See Table 6.3.) The main use of this analysis was to search for sex bias in the individual items. Item difficulty was computed as the proportion of correct answers on that item to the total number of responses; this analysis was done for boys, girls, and the combined group. Sex bias was evaluated for each item by computing a t ratio for the difference between the

Table 6.3

Problem Solving Item Analysis
Proportion of Correct Responses for Boys, Girls and Both

Item Number & Type	Boys	Girls	Both	t value of the Difference between Boys' and Girls' Proportions
1 (B)	.44	.51	.47	-.93
2 (A)	.11	.16	.13	-.46
3 (C)	.62	.45	.54	2.14*
4 (A)	.23	.22	.23	.32
5 (B)	.08	.14	.11	-1.02
6 (A)	.27	.31	.29	.19
7 (A)	.11	.13	.12	-.09
8 (B)	.21	.25	.23	-.46
9 (A)	.30	.27	.28	.84
10 (B)	.20	.23	.21	-.46
11 (C)	.39	.33	.36	.74
12 (C)	.62	.66	.64	-.65
13 (A)	.46	.37	.42	-1.02
14 (C)	.67	.61	.64	.74
15 (B)	.19	.14	.17	.74
16 (C)	.16	.11	.13	.93
17 (B)	.34	.43	.39	-1.30
18 (B)	.34	.29	.32	.83
19 (B)	.28	.19	.24	1.39
20 (C)	.27	.19	.23	1.29
21 (C)	.24	.36	.30	-.93
22 (C)	.38	.29	.34	1.11
23 (A)	.02	.01	.02	.0
24 (A)	.17	.20	.19	-.37

*p < .05

proportions, that is, on the difference between the item's difficulty for boys and for girls. (Downie & Heath, 1970) Only on Item 3 was there a difference significant at the .05 level. This item asks which player has the best shooting record given a table of shots attempted and shots made. It may well be that more boys than girls were familiar with that task; Item 19 which had the next largest (but nonsignificant) t ratio also involves sports. Item 3 was unsatisfactory for another reason; the correct answer could be obtained from the table by subtracting as well as by dividing. This was not discovered in the pilot test because Item 3 was one of the two new items used to replace other items found unsatisfactory in the pilot test. This item should not be used again for both reasons. On the whole, however, the test of mathematical problems seemed free of sex bias.

5. Sex-by-level ANOVAs. It was noticed that the boys and girls were not equally distributed between the two levels of mathematics classes. (See Table 6.4.) A χ^2 test showed that this distribution differed from the expected distribution at the .05 level. The facts that Level 1 students might have had more opportunity to learn the mathematics involved in the problem-solving test than Level 2 students and that a disproportionate number of girls were found in Level 1 might have obscured sex-related differences on the problems-solutions scales. The uneven distribution might also have affected the means on the

Table 6.4

Distribution of Boys and Girls
in Levels of Mathematics Classes

Level	Boys	Girls
1	43	51
2	50	32

spatial tests. In order to test these hypotheses, two-way, sex-by-level ANOVAs were performed on the scores from the five tests of visual spatial abilities and the solutions, diagrams, and correct representations subscales of the test of mathematical problem solving. The per cent scales were not included because their distributions were not normal and they did not seem to satisfy the assumptions for the analysis of variance. In all of the ANOVAs only those subjects with complete test data were used (170 of the original 176).

A set of exact ANOVAs with equal cell sizes of 32 was computed first; the cells were made equal by randomly eliminating subjects from the three cells with more than 32 subjects. This sample was accidentally destroyed after the analysis was finished. Since the author had intended to use an equal-cell-sized sample in a later analysis (a plan later abandoned), another sample

was randomly drawn with 32 in each cell and ANOVAs were computed using that data. There were some substantial differences between the two sets of F values. For this reason a set of ANOVAs with unequal cell sizes was computed using data from the 170 subjects who had taken all the tests. Table 6.5 presents the F values generated by these three sets of ANOVAs. Results for a test which were significant on at least two of the ANOVAs were considered important and are discussed in the following paragraphs. Others were suspected to be artifacts of the sample or to be due to limitations of this method of analysis.

In general the results of the three ANOVAs were the same as those generated by the t tests. On Form Board 2 boys' performance surpassed that of the girls at the .01 level; there was no significant sex-related difference on any of the other tests of visual spatial abilities. The ANOVAs showed significantly better performance by girls on Diagrams A and Diagrams T. The significant difference in favor of girls on Diagrams B computed using the t ratio was not reproduced by any of the ANOVAs. The significant differences in favor of girls on Correct Representations B and T obtained with the t ratio appeared only in the ANOVA using all the data.

In addition to confirming or denying the results of the t tests, the ANOVAs turned up some interesting sex-by-level interactions. The interactions between sex and level on

Table 6.5

F Values Generated by the Sex-by-Level ANOVAs

Tests	Equal Cell Sizes 1st Sample			Equal Cell Sizes 2nd Sample			Unequal Cell Sizes Using All Data		
	Sex	Level	Sex x Level	Sex	Level	Sex x Level	Sex	Level	Sex x Level
1. Form Board 2	4.63*	2.52	< 1	7.49**	< 1	6.62*	9.51**	< 1	3.01
2. Hidden Figures 2	< 1	2.12	2.12	< 1	5.08*	4.70*	< 1	5.46*	2.80
3. DAT Space Relations	< 1	8.45**	< 1	< 1	6.58*	< 1	< 1	9.98**	< 1
4. Card Rotations	< 1	9.33**	1.01	< 1	4.00*	< 1	< 1	5.78*	< 1
5. Cubes Comparison	2.20	10.48**	< 1	1.10	8.83**	< 1	2.48	10.68**	< 1
6. Problem Solutions A	< 1	16.60**	< 1	< 1	13.42**	< 1	< 1	22.91**	< 1
7. Problem Solutions B	1.45	39.35**	< 1	< 1	49.56**	3.56	< 1	52.74**	2.10
8. Problem Solutions C	8.77**	60.64**	< 1	6.28*	65.22**	< 1	6.53*	80.05**	< 1
9. Problem Solutions T	5.19*	64.81**	< 1	1.41	66.79**	< 1	2.78	86.01**	< 1
10. Diagrams A	5.17*	37.22**	5.72*	2.75	24.73**	3.14	4.11*	28.88**	3.33
11. Diagrams B	< 1	9.52**	2.67	2.05	11.59**	1.57	3.81	13.20**	3.50
12. Diagrams C	2.19	4.93*	< 1	1.09	3.02	1.09	1.84	5.33*	1.30
13. Diagrams T	5.09*	37.99**	6.78*	3.86	28.30**	3.86	6.39*	34.32**	5.36*
14. Correct Representations A	< 1	18.55**	3.40	1.74	16.88**	5.77*	1.25	22.07**	3.23
15. Correct Representations B	1.20	10.80**	< 1	2.59	13.55**	< 1	4.97*	16.21**	1.97
16. Correct Representations C	1.87	1.87	< 1	< 1	< 1	< 1	1.17	2.34	< 1
17. Correct Representations T	1.89	24.25**	3.36	3.51	23.55**	5.36*	3.99*	1.15	4.42*
			(df ₁ df ₂) = (1, 124)				(df ₁ , df ₂) = (1, 166)		

*p < .05

**p < .1

Diagrams T and Correct Representations T were significant ($p < .05$) for two of the three ANOVAs. Figure 6.2 is a graph of the cell means on Diagrams T for the ANOVA using all the data. It indicates that the sex-related difference on that scale was due to the superior performance of girls from the Level 1 classes. What makes this graph noteworthy is that the graphs of the cell means for all the scales involving diagrams or correct representations for all three ANOVAs are like this one, despite the fact that few of the interactions are significant at the .05 level.

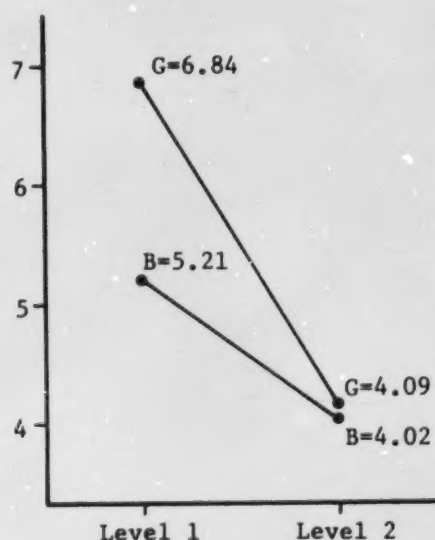


Figure 6.2. Sex-by-level interaction: cell means on Diagrams T ANOVA using all data.

6. Recomputation of the t test: Problem Solutions C. Another significant ($p < .05$) difference which appeared in all three ANOVAs. was one in favor of boys on the Problems Solutions C scale. This had been suggested by the results of the t test ($t = 1.13$), but the ratio was not significant at the .05 level. Since the one problem showing significant sex bias was a C-type problem, the author decided to recompute some of the statistics for the Problem Solutions C scale eliminating Problem 3. The means for boys and for girls became 2.73 and 2.55, respectively; the t ratio was reduced to .65.

7. Recomputation of the sex-by-level ANOVA: Problem Solutions C. The effect on the F value was investigated by recomputing the ANOVA using the second sample with equal cell sizes. Removing the sex-biased item reduced the F value for a main effect due to sex from 6.28 ($.01 < p < .025$) to 2.58 ($p > .05$). These decreases in the sex-related differences on the Problem Solutions C scale demonstrate the importance of checking test items for sex bias.

8. Sex-by-IQ ANOVAs. Just as the distribution of the boys and girls in the levels was uneven, the distribution of boys and girls on the IQ stanines appeared to be uneven. (See Table 6.6.) The stanines were divided into three levels: Stanines 7 through 9 were Group 1, Stanines 5 and 6 were Group 2, and Stanines 2 through 4 were Group 3. A χ^2 test showed that the distribution

Table 6.6
Distribution of Boys and Girls in IQ Stanines

Stanine	Boys		Girls
9	8		4
8	12		14
7	<u>13</u>	Group 1	<u>16</u>
	33		34
6	15		20
5	<u>17</u>	Group 2	<u>14</u>
	32		34
4	18		7
3	5		3
2	<u>3</u>	Group 3	<u>1</u>
	26		11

IQ scores were not available for 2 boys and 4 girls.

among groups was only marginally different from that expected ($.05 < p < .10$). Nevertheless, the author decided to examine the data separated into these three IQ groups using analysis of variance as was done with level of mathematics class. In order to perform ANOVAs with equal cell sizes of 11, subjects were randomly eliminated from five of the six cells. Since the small number of subjects in each cell strained the assumptions

necessary for analysis of variance, the F ratios resulting from these ANOVAs which are given in Table 6.7 should be interpreted with caution.

In the sex-by-IQ ANOVAs significant sex-related differences on Form Board 2 in favor of boys and on Diagrams T in favor of girls emerged as in all the other analyses. The significant differences in favor of girls on scales involving diagrams which had been demonstrated in the previous analyses failed to appear. The one exception was on Diagrams B on which the t-test results had shown girls to be significantly better. Boys also performed significantly better than girls on Cubes Comparison in this subsample which was suggested by the t-test results but was not significant.

The surprising result of the sex-by-IQ ANOVAs was the appearance of a sex-related difference in favor of boys on DAT Space Relations ($p < .01$). This difference was suspected to be an artifact of the subsample which was only 66 of the original 176. However, since the norms for boys and for girls reported in the manual for the DAT differed more than the means for the whole group in this study, it seemed important to pursue this investigation. Cell means were computed for boys and girls in the three IQ groups using scores from the whole group of 170. These cell means and those from the sex-by-IQ ANOVA are presented for comparison in Table 6.8. It appears that

Table 6.7

F Values Generated by the Sex x IQ ANOVAs (Equal Cell Sizes)

Tests	Sex	IQ	Sex x IQ
1. Form Board 2	7.31**	3.22*	< 1
2. Hidden Figures 2	1.05	1.77	< 1
3. DAT Space Relations	10.27**	10.98**	1.24
4. Card Rotations	< 1	1.77	< 1
5. Cubes Comparison	6.46*	14.44**	< 1
6. Problem Solutions A	< 1	11.27**	< 1
7. Problem Solutions B	< 1	28.76**	1.00
8. Problem Solutions C	3.55	28.47**	< 1
9. Problem Solutions T	2.01	40.48**	< 1
10. Diagrams A	1.73	8.33**	< 1
11. Diagrams B	4.59*	5.50**	< 1
12. Diagrams C	2.86	< 1	< 1
13. Diagrams T	4.77*	10.05**	< 1
14. Correct Representations A	1.91	8.37**	< 1
15. Correct Representations B	< 1	5.88**	< 1
16. Correct Representations C	1.00	18.42**	< 1
17. Correct Representations T	2.56	11.60**	< 1
$(df_1, df_2) = (1, 60)$ $(df_1, df_2) = (2, 60)$ $(df_1, df_2) = (2, 60)$			

*p < .05

**p < .01

the subsample did perform differently on DAT Space Relations than the whole group; in all cells but the girls' IQ Group 3 (in which the sample and the whole group were the same) the scores of boys in the subsample were higher than those of boys in the whole group, and scores of girls in the subsample were lower than those of girls in the whole group. The sex-related differences found in the IQ subsample were not found in the group as a whole, except in the lowest IQ group.

Table 6.8

Cell Means for the DAT Space Relations Test
Comparison of the Sex-by-IQ Subsample and the Whole Group

IQ Groups	Girls		Boys	
	Whole Group	IQ Subsample	Whole Group	IQ Subsample
Stanines 7-9	29.9	27.6	30.3	35.4
Stanines 5-6	26.6	23.9	23.9	26.1
Stanines 2-4	14.6	14.6	23.7	24.6

The significance of the sex-related difference in the Cubes Comparison scores was then also suspected to be a result of the sample. Cell means for the three IQ groups were computed using the 170 Cubes Comparison scores. Table 6.9 presents these and the cell means from the sex-by-IQ ANOVA. The differences between

the two sets of means were not as pronounced as those in the preceding comparison. However, again there seemed to be a substantial difference at the lowest IQ level.

Table 6.9

Cell Means for the Cubes Comparison Test
Comparison of the Sex-by-IQ Subsample and the Whole Group

IQ Groups	Girls		Boys	
	Whole Group	IQ Subsample	Whole Group	IQ Subsample
Stanines 7-9	29.9	30.7	33.0	35.5
Stanines 5-6	26.3	25.0	26.3	27.7
Stanines 2-4	20.6	20.6	26.0	25.2

The analyses of this section can be summarized best by returning to the hypotheses stated at the beginning of this discussion. Boys and girls in this sample did not differ significantly in their ability to solve mathematical problems of any of the three types, as was hypothesized. The sex-related difference on the Problem Solutions C scale, significant only in the sex-by-level ANOVA, was in fact due to sex bias in one item. The results from the spatial tests only partly supported the second hypothesis. Significant sex-related differences in favor of boys on Form Board 2 were found in all analyses. A significant difference in favor of boys was observed on Cubes Comparison in the sex-by-IQ

ANOVA. This difference had been suggested by the results of the t test but was not significant; it was most pronounced at the lowest IQ level. The sex-by-IQ ANOVA of the DAT Space Relations scores also showed a significant main effect for sex. Examination of the cell means indicated that this difference was probably an artifact of the sample. The results pertaining to the first part of the fifth hypothesis were also mixed. Girls drew or modified significantly more diagrams than boys, especially for A-type problems, and then correctly represented more problems although these differences seemed to be confined mostly to subjects in the Level 1 mathematics classes.

Part B. Relationships Between Visual Spatial Abilities
and Mathematical Problem Solving Ability and
Sex-Related Differences in These Relationships

The purpose of the last section was to search for sex-related differences in visual spatial abilities and mathematical problem-solving ability. In this section the relationships between the spatial and mathematical variables are examined. The question of whether or not the relationships are the same for boys and for girls is also discussed. Specifically the third hypothesis, the fourth hypothesis, and the last part of the fifth hypothesis are investigated.

3. There are significant positive relationships between each of the visual spatial abilities and mathematical

problem solving ability. The relationships are stronger for Vz tests than for SR-O tests and stronger for three-dimensional tests than for two-dimensional tests. These relationships do not differ by sex.

4. Each type of visual spatial ability is more closely related to solving mathematical problems with high spatial content than to solving problems with little spatial content.
5. Use of a diagram in solving a mathematical problem is positively related both to the ability to solve that problem and to visual spatial abilities. There are no sex-related differences in these relationships.

In order to examine these relationships from several different points of view, a number of statistical analyses were performed. These included computation of correlation coefficients, factor analysis, regression analysis, and canonical correlation analysis, all done separately by sex. Figure 6.3 indicates the sequence of these analyses.

1. Correlation coefficients. First correlation coefficients were computed for the tests of visual spatial abilities, the problem-solving subscales, and IQ using the 170 subjects who had taken all the tests. The matrices of these coefficients, which were computed for boys, for girls, and for the total group, are found in Appendix E. Since significance of a correlation coefficient is related to the number of subjects or observations,

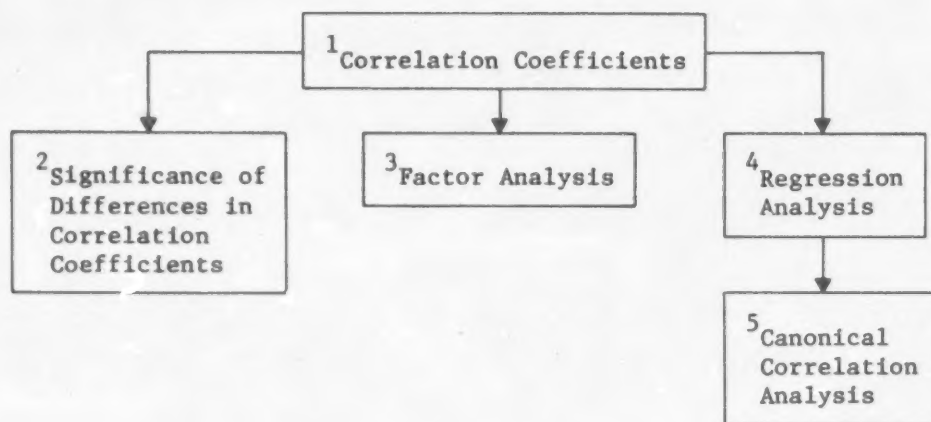


Figure 6.3. Sequence of data analysis, part B.

the .05 and .01 levels of significance were slightly different for boys and for girls. However, using the upper bound for each set (.22 for the .05 level and .29 for the .01 level) would cause little distortion in interpretation. The significance levels for the total group were lower: .15 for the .05 level and .20 for the .01 level.

As can be seen by examining the correlation matrices with these significance levels in mind, a number of the coefficients

achieved the .05 or .01 level of significance. For girls and for the whole group, the relationships between all the spatial tests and all the problem-solutions scales were significant and positive. For boys the DAT Space Relations test was significantly related to all the problem-solutions scales. Problem Solutions A was not significantly related to Cubes Comparison or Card Rotations for boys, although the relationships of the other problem-solutions scales to these two spatial tests were significant. Form Board 2 and Hidden Figures 2 were significantly related to none of the problem-solutions scales for boys.

The correlation matrices also provided information about the relationships between use of diagrams and both mathematical problem solving ability and visual spatial abilities. For the girls and the whole group, all of the spatial scales except Form Board 2 were significantly related to correctly representing A-type problems; all except Form Board 2 and Hidden Figures 2 were also related to correctly representing B-type problems. The spatial tests, except Form Board 2, were also significantly related to the total number of correct representations. For boys only the DAT Space Relations test and Card Rotations were significantly related to correctly representing problems with diagrams. The relationships between correctly representing a problem with a diagram and solving that problem are summarized in Table 6.10. As was expected, the size of the correlation

coefficients was directly related to the type of problem: use of a diagram was very important in A-type problems, less important in B-type problems, and unimportant in C-type problems.

Table 6.10

The Relative Importance of Using Diagrams
in Solving Three Types of Problems

	Girls	Boys	Both
<u>Problem Solutions A</u>			
Diagrams A	.362	.035	.326
Correct Representations A	.702	.662	.669
Per Cent A	.684	.655	.667
<u>Problem Solutions B</u>			
Diagrams B	.359	.109	.234
Correct Representations B	.538	.235	.399
Per Cent B	.384	.172	.279
<u>Problem Solutions C</u>			
Diagrams C	.055	-.011	.017
Correct Representations C	.108	-.066	.028
Per Cent C	.108	-.066	.028

2. Significance of differences in correlation coefficients.

It appeared that the correlation coefficients were generally larger for girls than for boys especially between the spatial and mathematical variables. In order to investigate the significance of the differences between boys' and girls' correlation coefficients, a Fisher r to Z transformation was performed on each coefficient and the differences tested for significance using a t ratio (Downie & Heath, 1970). Table 6.11 presents the results of this investigation.

A total of 231 pairs of correlation coefficients were tested for significance of their differences. One could expect that by chance 12 of the differences would be significant at the .05 level and 3 at the .01 level. The actual results were that 17 of the differences were significant at the .05 level; of these 8 were significant at the .01 level. While the actual results were not much above the chance expectation, the pattern of results is interesting. All but 3 of the 17 correlations on which boys and girls differed were involved with diagrams for C-Type problems which supposedly have no spatial content. For boys this was moderately related to high scores on Form Board 2 or Card Rotations but not at all to the other problem scales. For girls, drawing correct diagrams for C-Type problems was moderately related to some of the other solution and diagram scales but not at all to Form Board 2. It should also be noted

Table 6.11

Significant Sex-Related Differences in Correlation Coefficients

Variables	1	2	3	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. IQ																							
2. Form Board 2																							
3. Hidden Figures 2																							
4. DAT Space Relations																							
5. Problem Solutions A																							
6. Problem Solutions B																							
7. Problem Solutions C																							
8. Problem Solutions T																							
9. Diagrams A																							
10. Diagrams B																							
11. Diagrams C			B** ^a							G*													
12. Diagrams T																							
13. Correct Representations A																							
14. Correct Representations B																							
15. Correct Representations C			B**											G*									
16. Correct Representations T												G*				G*							
17. Per Cent A																							
18. Per Cent B																							
19. Per Cent C			B**			G*		G**		G**								G**					
20. Per Cent T								G*	G*												G**		
21. Card Rotations																					B**		
22. Cubes Comparison																	G*				G**		

^aB = The boys' correlation coefficient between these two variables was significantly higher than the girls'.

*p < .05

**p < .01

that Form Board 2 is the one spatial test on which sex-related differences in performance have consistently appeared in this sample. (See the preceding section of this chapter.)

The test for significance of differences in correlation coefficients appears to be rather stringent: the difference had to be about .30 in order to achieve the .05 level of significance. There may be other sex-related differences in the relationships that this analysis was not sensitive enough to detect. For the present, however, it seems that there are few sex-related differences in the relationships between the spatial and mathematical variables or between the diagrams and solutions scales of the test of mathematical problem solving.

The relative importance of Vz and Sx-0 tests or two- and three-dimensional tests in predicting mathematical problem solving ability can also be examined using coefficients from the correlation matrices. Table 6.12 contains the correlation coefficients between Problem Solutions Total and the four spatial tests for the whole group of boys and girls. The three-dimensional tests in both columns were somewhat better predictors of the total problem-solutions scale, but the difference was significant only for the Vz test. No such generalization can be made across the rows in the table. The same patterns can be observed in the correlation

coefficients for the problem-solutions subscales, although the differences were sometimes less significant.

Table 6.12

Correlations Between the Spatial Variables
and Problem Solutions Total for the Whole Group

	SR-O	Vz	<u>t</u> ratio ^a (SR-O-Vz)
2-Dimensional	.350	.234	-1.41
3-Dimensional	.395	.446	.76
<u>t</u> ratio (3-D-2-D).57		2.72*	

^a t ratio for the significance of the difference between correlation coefficients, correlated data. (Downie & Heath, 1970)

*
p < .01

3. Factor Analysis. Although the correlation matrices provide information related to the hypotheses under investigation in this section, there are limitations to this method of picking and choosing among the coefficients, especially the possibility of not seeing the forest for the trees. In order to achieve some reduction of the data, a series of factor analyses were performed. This study had not been planned to be primarily factor analytic, but the sex-related differences in factor

structure noted in Chapters II and III indicated that factor analysis might provide interesting results. Because the factor analysis was considered exploratory, commitment to one type of analysis seemed unwise. The method used in the CAA Project (Harris & Harris, 1973) and discussed in Chapter II suggested the possibility of using several different types of factor analysis, each with orthogonal and oblique rotations, and then finding factors common among them. The per cent scales and total scales, which were algebraic combinations of some of the other scales, which were not included in the factor analysis.

Principal Components Analysis (a noncommunality solution) and Principal Factor Analysis (a communality solution) were the two methods of factor analysis used. The Principal Components Analysis (P COMP A) was programmed to stop extracting factors when the eigen values reached one; in the Principal Factor Analysis (P FACT A) they were allowed to go to zero. This meant that more factors were extracted by the latter method than by the former. Each method was done on the correlation matrices for boys and girls separately, first with an orthogonal rotation and then with two oblique rotations. The oblique rotations selected were Harris-Kaiser Procedures, Types 1 and 2.

Oblique rotations, although theoretically less satisfying, are often used to clarify the factors by eliminating tests with low loadings. The Type I oblique rotation was not particularly useful in that respect, and in fact appeared to complicate some

of the factors. For example, the Type I rotation of the P COMP A of the boys' data added Diagrams C and Correct Representations B and C to the factor containing the spatial tests. Since the loss in theoretical justification did not seem to be balanced by any gain in clarity, the results of the Type I oblique rotation are not discussed further.

Factors which emerged common to the orthogonal and oblique (Type 2) rotations of P COMP A and P FACT A are presented in Tables 6.13 through 6.16. The order of the factors is arbitrary as they all accounted for about the same amount of the variance and the order of emergence differed from analysis to analysis. Variables were listed under a factor if their loading on that factor exceeded .30 for any of the analyses for either sex. Four factors emerged consistently in all four analyses: a Solutions factor, a Space factor, and two Drawing factors, one for A- and B-type problems and the other for C-type problems.

The high loadings of Problem Solutions A, B, and C on the first factor indicate that this should be called the Solutions factor. (See Table 6.13.) In the P COMP A, Diagrams A and Correct Representations A also had significant loadings on the Solutions factor. Since the A-type problems were defined as those in which the stimulus is pictorial or which require spatial or geometric knowledge for solution, it was expected that the correct pictorial representations of A-type problems

Table 6.13
The Solutions Factor

Variables	Principal Components Analysis				Principal Factor Analysis							
	Orthogonal		Oblique		Orthogonal				Oblique			
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
. Problem Solutions A	72 ^a	86	70	84	57	43	37	68	49	41	29	62
. Problem Solutions B	39	71	37	65	38		63	35	37		54	
. Problem Solutions C	60	62	60	54	59		72		62		72	
. Hidden Figures 2	72	26	78	32								
. Diagrams A	40	50	37	26		67		53		17		38
. Correct Representations A	72	78	65	65	27	75	11	82		42		81

Decimal points omitted

would be closely related to their solutions. There were two noticeable differences in the boys' and girls' loadings in the P COMP A. The girls' loading of Problem Solutions B was smaller than the boys' loading of that test. (As can be seen in Table 6.14 Problem Solutions B has a higher loading on the Space factors for girls than for boys.) The other difference is that the loading of Hidden Figures 2 on the Solutions factor is higher for girls than for boys. This could be an indication that boys and girls use different problem-solving methods on Hidden Figures items, perhaps that boys use a more visual method and girls a more logical method.

In the P FACT A the Solutions factor split into two sub-factors, one concerned with solutions alone and the other with A-type problems. This further reinforces the idea that pictorial representations are essential to A-type problems. The factor loadings of Problem Solutions A on the two subfactors suggest that the pictorial representation may be more important for boys than for girls. However, checking the entries in the matrices of correlation coefficients did not confirm this hypothesis. In the P PACT A which allowed more factors, Hidden Figures 2 had its own factor.

A single Space factor emerged in the P COMP A with all the spatial tests except Hidden Figures 2 loading heavily on it. (See Table 6.14.) Hidden Figures 2 did not even appear

Table 6.14
The Space Factor

Variables	Principal Components Analysis				Principal Factor Analysis					
	Orthogonal		Oblique		Orthogonal		Oblique			
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
1. Cubes Comparison	75 ^a	77	75	76	61	61	42	13	61	08
2. DAT Space Relations	69	74	67	71	56	69	35	13	51	21
3. Form Board 2	71	59	71	56	58	52	13	45	19	36
4. Card Rotations	64	71	60	70	55	54	03	33	32	18
5. Problem Solutions B	52	41	41	31	48	35	19	21	20	20
6. Problem Solutions A	38	06			29	14				
7. Problem Solutions C	34	28			27	17				
8. Correct Representation A	30	01			22	13				

^aDecimal points omitted

on this factor. Problem Solutions B had a substantial loading on the Space factor for girls and a somewhat more modest loading for boys. In addition, the other two problem solutions scales had modest loadings on the Space factor for girls but these disappeared with the oblique rotation. In the P FACT A with orthogonal rotation there was only one Space factor, and it was very similar to the one obtained by the P COMP A. With oblique rotation, however, two subfactors emerged. The subdivision for girls is clearly along dimensional lines while the boys' subdivision might be dimensional except for the loadings of Card Rotations.

The first Drawing factor was for A- and B-type problems. (See Table 6.15.) It included Diagrams A and B as well as Correct Representations A and B. The loadings for Correct Representation A were lower than the others in all the analyses, probably because that test had such high loadings on the Solutions factor. In the P FACT A with oblique rotation, the girls' factor split neatly into two subfactors, one for A-type problems and one for B-type problems. The appearance of Problem Solutions B on this factor for girls but not for boys again suggests that B-type problems were more closely linked to spatial skills for girls in this sample. However, it should be noted that this loading disappeared with oblique rotation.

The fourth factor which emerged was a Drawing factor for C-type problems, those which appeared to have no spatial content.

Table 6.15

The First Drawing Factor (Problem Types A and B)

Variables	Principal Components Analysis				Principal Factor Analysis			
	Orthogonal		Oblique		Orthogonal		Oblique	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
1. Diagrams A	55 ^a	65	46	71	37	53	63	48
2. Diagrams B	86	85	87	87	65	76	24	72
3. Correct Representations A	34	29	19	41	20	22	39	19
4. Correct Representations B	81	87	79	88	69	77	11	80
5. Problem Solutions B	46	06			46	10		

^aDecimal points omitted

(See Table 6.16.) Form Board 2 had modest loadings on this factor for boys but not at all for girls. It should be remembered that this was one of the few sex-related differences between correlation coefficients that was significant at the .05 level.

The results of the factor analyses can be summarized as follows. Solving B-type problems seems to be more closely related to spatial skills than are the other problem-solving subscales; the relationship appears to be stronger for girls than for boys in this sample. Hidden Figures 2 had substantial loadings on the Solutions factor for girls but not for boys. Form Board 2 had modest loadings on the second Drawing factor (C-type problems) for boys but not for girls. With those two exceptions, the spatial tests confined themselves to the Space factor. There was some evidence of splitting of the Space factor but it seemed to be more along dimensional lines than to be an SR-O, Vz split.

4. Regression analysis. In order to investigate the relative importance of each of the spatial tests in predicting scores on the problem-solving scales, two regression analyses were carried out, each for boys and for girls. The first regression analysis used all the spatial tests including Hidden Figures 2; in the second analysis Hidden Figures 2 was omitted since it had not appeared on the Space factor. The order of entering the spatial variables was set by the experimenter and was based on the relative sizes of the correlation coefficients for the total group

Table 6.16

The Second Drawing Factor (Problem Type C)

Variables	Principal Components Analysis				Principal Factor Analysis			
	Orthogonal		Oblique		Orthogonal		Oblique	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Diagrams C	91 ^a	94	92	95	87	89	86	88
Correct Representations C	93	91	94	92	86	90	85	85
Form Board 2	- 16	36	- 18	36	- 09	25	- 08	37
Hidden Figures 2	- 09	31			- 02	11		

^aDecimal points omitted.

between the spatial tests and the particular problem-solving scale being predicted. Table 6.17 lists the standardized regression coefficients for both regression analyses for boys and for girls. Standardized coefficients were used to facilitate comparison among the independent variables (the spatial tests). The coefficients of determination were also included in Table 6.17 to show how much of the variance in each problem-solving scale was accounted for by these "best" linear combinations of the spatial scores.

Examination of the parts of Table 6.17 concerned with the four problem-solutions scales leads to several generalizations. In each case the regression equation accounted for more of the variance on the problem scale for girls and than for boys. The regression equations for Problem Solutions B predicted more of its variance than did the equations for Problems Solutions A or C. Both these results agree with the results of the factor analyses. Hidden Figures 2 contributed less to the regression equations for boys than for girls in all four cases, but little predictability was lost for either sex by excluding it. In the equations for Problem Solutions A and B, DAT Space Relations was the most important predictor for boys; for girls all but Form Board 2 were important predictors. In the equations for Problem Solutions C, however, the DAT was the most important predictor for girls while Cubes Comparison was important for boys. The equation for Problem Solutions T combined the trends

Table 6.17
Standardized Regression Coefficients

Variables	Problem Solutions A			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Hidden Figures 2	.18*	.11		
DAT Space Relations	.19	.23*	.20*	.23*
Cubes Comparison	.23*	-.04	.24**	-.04
Card Rotations	.16	.06	.18	.08
Form Board 2	-.02	.07	.00	.07
Coefficient of Determination	.26	.10	.23	.09

Variables	Problem Solutions B			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Hidden Figures 2	.18**	.09		
DAT Space Relations	.22**	.32***	.23**	.32***
Cubes Comparison	.28**	.10	.29***	.11
Card Rotations	.17	.17	.20*	.18*
Form Board 2	.04	-.02	.06	-.03
Coefficient of Determination	.38	.24	.35	.23

*p < .10 **p < .05 ***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Problem Solutions C			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Hidden Figures 2	.21**	.03		
DAT Space Relations	.35***	.05	.36***	.05
Cubes Comparison	.05	.26**	.05	.26**
Card Rotations	.02	.15	.05	.15
Form Board 2	.05	-.07	.07	-.07
Coefficient of Determination	.24	.13	.20	.13

Variables	Problem Solutions T			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Hidden Figures 2	.23**	.09		
DAT Space Relations	.32***	.23*	.33***	.23*
Cubes Comparison	.22*	.15	.22**	.15
Cards Rotations	.14	.16	.17	.17
Form Board 2	.03	-.02	.06	-.02
Coefficient of Determination	.41	.19	.36	.18

*p < .10 **p < .05 ***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Diagrams A			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.32**	.19	.32**	.19
Cubes Comparison	.16	.10	.16	.11
Card Rotations	.03	.07	.04	.08
Hidden Figures 2	.11	.04		
Form Board 2	-.09	-.16	-.07	-.16
Coefficient of Determination	.19	.08	.18	.08

Variables	Diagrams B			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.27**	.20	.27**	.20
Card Rotations	-.08	.19	-.07	.20
Form Board 2	.09	.07	.09	.07
Hidden Figures 2	.05	.02		
Cubes Comparison	-.06	-.22*	-.06	-.21*
Coefficient of Determination	.07	.10	.07	.10

*p < .10

**p < .05

***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Diagrams C			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Hidden Figures 2	.07	.19*		
DAT Space Relations	.21	-.09	.21	-.09
Form Board 2	-.07	.29**	-.06	.29**
Cubes Comparison	.04	.03	.04	.04
Card Rotations	-.24*	.00	-.23*	.04
Coefficient of Determination	.08	.12	.07	.08

Variables	Diagrams T			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.37***	.22	.38***	.22
Cubes Comparison	.11	-.02	.11	-.02
Card Rotations	-.03	.14	-.02	.15
Hidden Figures 2	.11	.05		
Form Board 2	-.04	-.05	-.03	-.05
Coefficient of Determination	.18	.08	.17	.08

*p < .10

**p < .05

***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Correct Representations A			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.25**	.34**	.26**	.34**
Cubes Comparison	.25**	-.05	.25**	-.05
Card Rotations	.09	.07	.12	.08
Hidden Figures 2	.24**	.02		
Form Board 2	-.04	-.09	-.01	-.09
Coefficient of Determination	.30	.11	.25	.11

Variables	Correct Representations B			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Card Rotations	.14	.34***	.15	.32***
DAT Space Relations	.22*	.08	.22*	.08
Cubes Comparison	.12	-.05	.12	-.05
Hidden Figures 2	.03	-.10		
Form Board 2	-.05	-.24**	-.05	-.24**
Coefficient of Determination	.13	.13	.13	.11

*p < .10

**p < .05

***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Correct Representations C			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.21	.02	.21	-.02
Cubes Comparison	.18	-.06	.19	-.05
Hidden Figures 2	.08	.10		
Card Rotations	-.20	-.03	-.19	-.01
Form Board 2	-.20	.28**	-.20	.26**
Coefficient of Determination	.12	.09	.11	.08

Variables	Correct Representations T			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.27**	.30**	.27**	.30**
Cubes Comparison	.27**	-.08	.27**	-.08
Card Rotation	.10	.21*	.13	.20*
Hidden Figures 2	.20**	-.04		
Form Board 2	-.08	-.14	-.06	-.14
Coefficient of Determination	.30	.13	.26	.13

*p < .10 **p < .05 ***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Per Cent A			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.15	.28**	.15	.28**
Card Rotations	.15	.07	.17	.07
Cubes Comparison	.13	-.07	.13	-.06
Hidden Figures 2	.18*	-.02		
Form Board 2	.01	.02	.02	.02
Coefficient of Determination	.17	.09	.14	.09

Variables	Per Cent B			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Card Rotations	.19	.28**	.20	.25**
Cubes Comparison	.07	.10	.07	.10
DAT Space Relations	.13	-.06	.13	-.06
Hidden Figures 2	.05	-.17		
Form Board 2	-.09	-.23*	-.09	-.22*
Coefficient of Determination	.08	.11	.08	.09

*p < .10

**p < .05

***p < .01

Table 6.17 (Continued)
Standardized Regression Coefficients

Variables	Per Cent C			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
Cubes Comparison	.18	-.06	.19	-.05
DAT Space Relations	.21	.02	.21	.02
Card Rotations	-.20	-.03	-.19	-.01
Hidden Figures 2	.08	.09		
Form Board 2	-.20	.28**	-.20	.28**
Coefficient of Determination	.12	.09	.11	.08

Variables	Per Cent T			
	Regression 1		Regression 2	
	Girls	Boys	Girls	Boys
DAT Space Relations	.21*	.11	.21*	.11
Card Rotations	.15	.19	.17	.15
Cubes Comparison	.14	-.04	.14	-.06
Hidden Figures 2	.16	-.20*		
Form Board 2	-.04	.02	-.02	.03
Coefficient of Determination	.18	.08	.16	.04

*p < .10 **p < .05 ***p < .01

of the three subscales but DAT Space Relations still accounted for more of the variance than the other tests.

Fewer sex-related differences are observable in the regression analyses for the scales involved with drawing a diagram. The spatial variables accounted for a little more of the variance in Diagrams A for girls than for boys but for the other diagrams scales there was little difference. Overall DAT Space Relations was the best predictor of the diagrams scales, and it appears to predict a little better for girls than for boys.

The scales for correct representation and per cent can be discussed together because the same patterns are observable in both. Except for the scales for A-type problems and the totals (and then only for girls), the regression equations accounted for only about 10% of the variance. For the girls the weight of prediction was equally split among several spatial variables. For the boys one spatial test usually carried most of the weight: for Correct Representations A and T, it was DAT Space Relations; for Correct Representations B, it was Card Rotations; and for Correct Representations C, it was Form Board 2. The coefficients of determination for solving A-type problems are very similar. This was expected given the definition of an A-type problem. What is surprising is that although the spatial variables accounted for a substantial amount of the variance in solving B-type problems, they did not predict much of the Correct Representations B scale.

Preliminary analysis using comparison of correlation coefficients had indicated that the three-dimensional tests were better predictors of the problem-solutions scales than the two-dimensional tests. (See Table 6.13.) Regression analysis was used to investigate this hypothesis further. Table 6.18 gives the standardized regression coefficients and coefficients of determination for predicting the problem-solution scales from the two- and three-dimensional tests separately. By comparing their coefficients of determination, one can see that the three-dimensional tests accounted for more of the variance on each problem-solution scale than did the two-dimensional tests. The differences ranged from 3% to 15% of the variance. The differences were larger for the girls than the boys.

5. Canonical correlation analysis. In order to obtain a clearer picture of the relationship between the spatial variables and a composite for each type of problem-solving scale, a canonical correlation analysis was run on the boys' and girls' data separately. Since Hidden Figures 2 had not added much to the prediction in the regression equations, it was not used in this analysis. The spatial tests were used as independent variables in each analysis. The three subscores for problem solutions, for diagrams, for correct representations, and for per cents were used as dependent variables in four different analyses.

Table 6.18
Standardized Regression Coefficients
Two- and Three-Dimensional Comparison

Variables	Problem Solutions A			Problem Solutions B		
	Girls	Boys	Both	Girls	Boys	Both
Two-Dimensional						
Form Board 2	.12	.13	.12	.21	.08	.11
Card Rotations	.28	.15	.22	.32	.33	.33
Coefficients of Determination	.12	.05	.07	.19	.13	.15
Three-Dimensional						
DAT Space Relations	.24	.28	.27	.29	.37	.33
Cubes Comparison	.28	-.01	.19	.34	.14	.24
Coefficient of Determination	.20	.08	.14	.30	.20	.25
Variables	Problem Solutions C			Problem Solutions T		
	Girls	Boys	Both	Girls	Boys	Both
Two-Dimensional						
Form Board 2	.15	-.02	.09	.21	.07	.10
Card Rotations	.17	.25	.21	.30	.30	.32
Coefficient of Determination	.07	.06	.06	.18	.11	.16
Three-Dimensional						
DAT Space Relations	.39	.08	.23	.38	.28	.33
Cubes Comparison	.08	.28	.19	.27	.18	.23
Coefficient of Determination	.19	.13	.11	.33	.16	.24

The only canonical correlation on which more than 5% of the variance in the dependent variables was accounted for was the first one for the problem solutions composite. The standardized coefficients are reported in Table 6.19. This canonical correlation is illustrative because it summarizes a number of results of the regression analysis. The canonical correlation accounted for a larger percentage of the girls' variance than of the boys' variance. In the girls' correlation scores for both A- and B-type problems were related to that part of the variance predicted by the spatial variables; in the boys' correlation only the scores for the B-type problems were important. In the girls' correlation, three of the spatial tests shared the prediction of the problem-solving composite. In the boys' correlation, the DAT Space Relations tests was most important. Since the coefficients for Card Rotations are almost the same for both sexes, this might be stated another way: for girls Cubes Comparison and DAT Space Relations have equal weight, but for boys the DAT took the greater share.

In summary, it can be said that most of the hypotheses concerning the existence of relationships were supported. There were, however, a few exceptions. Some of the spatial variables were not significantly related to problem solving ability for boys. In general, the spatial scores were more closely related to the scales from the problem solving test for

Table 6.19
 Canonical Correlation Analysis
 Spatial Composite with Problem Solutions Composite

Variables	Standardized Coefficients	
	Girls	Boys
Dependent		
Problem Solutions A	.39	.06
Problem Solutions B	.72	.99
Problem Solutions C	.05	-.03
Independent		
Form Board 2	.08	-.04
DAT Space Relations	.41	.67
Card Rotations	.34	.38
Cubes Comparison	.48	.20
Amount of the Variance in the Dependent Variables Accounted for by this Canonical Correlation	13.31%	7.78%

girls than for boys. With the exception of Correct Representations A (and then only for girls), the scales involved with diagrams were less closely related to the spatial variables than were the solutions scales. The hypothesis that the Vz tests would be better predictors of the problem-solutions scales than the SR-O tests was not supported, although the dimensional division did seem important. Another exception was that the spatial variables were more closely related to solving B-type problems than A-type problems. Interpretation of these results and an attempt to integrate them with the results of other studies are found in the final chapter.

Technical Report No. 387
(Part 3 of 3 Parts)

THE INTERRELATIONSHIP OF SEX, VISUAL SPATIAL ABILITIES,
AND MATHEMATICAL PROBLEM SOLVING ABILITY IN GRADE SEVEN

by

Ann Koch Schonberger

Report from the Project on Conditions of School
Learning and Instructional Strategies

Thomas A. Romberg
Faculty Associate

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

June 1976

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in the University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

WISCONSIN RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING

MISSION

The mission of the Wisconsin Research and Development Center for Cognitive Learning is to help learners develop as rapidly and effectively as possible their potential as human beings and as contributing members of society. The R&D Center is striving to fulfill this goal by

- conducting research to discover more about how children learn
- developing improved instructional strategies, processes and materials for school administrators, teachers, and children, and
- offering assistance to educators and citizens which will help transfer the outcomes of research and development into practice

PROGRAM

The activities of the Wisconsin R&D Center are organized around one unifying theme, Individually Guided Education.

FUNDING

The Wisconsin R&D Center is supported with funds from the National Institute of Education; the Bureau of Education for the Handicapped, U.S. Office of Education; and the University of Wisconsin.

ACKNOWLEDGEMENTS

Academic scholarship is not always the lonely pursuit it is often thought to be. Writing this acknowledgement provided me with an opportunity to recall the many pleasant relationships established during the research and writing of this dissertation.

First, I wish to thank my dissertation committee. Dr. Thomas Romberg, my major professor, first interested me in mathematical problem solving and assisted me greatly in the design, measurement, and data analysis aspects of my study. Dr. Elizabeth Fennema, whose interest in the effects of sex and spatial ability on mathematics achievement inspired my own interest, provided detailed criticism of this manuscript. Dr. Ann Becker contributed an important outside perspective to the study. I am grateful to all of them for their personal, as well as intellectual, support. Dr. Romberg allowed me to complete this study and other degree requirements in a rather unconventional way. Dr. Fennema and Dr. Becker provided examples of academic careers compatible with motherhood. Dr. Thomas Carpenter, Dr. John G. Harvey, and Dr. Mary Montgomery also read the manuscript and offered helpful comments.

Exchanging ideas about mathematical problem solving with other graduate students was useful to me in formulating my own ideas. The work of Ruth Ann Meyer, Diana Wearne, Don Whitaker, and Don Zalewski

was particularly important. In addition, Diana Wearne and Clement Steele took time from their own studies to code some of the problem solving tests.

The staff of the Bangor Public Schools cooperated by providing subjects for the study. Superintendent Wendell Eaton gave permission for the study. Mr. George Voss, Director of Research and Evaluation, made the necessary arrangements and provided some of the data. Mr. Lionel Lord and Mrs. Nancy Luttrell helped with the testing at the Fifth Street School. Thanks also go to the principals, the mathematics teachers, and the seventh-grade students at the Fifth Street and Union Street Junior High Schools who gave their time and cooperation.

I would also like to thank those who assisted me in the data analysis phase of my study. Dr. Jacob Evanson helped plan the analysis and Ms. Linda Junker assisted in the actual programming. The key-punching and computer operating staff of the Wisconsin Research and Development Center for Cognitive Learning provided competent and efficient service. Mrs. Dorothy Egner, also on the Center's staff, did the final typing.

Finally, I acknowledge the contributions of my family: my parents, my husband, and my children. My parents encouraged my intellectual activities and provided me with a first-class education. My children tolerated my absence in good humor. My husband unselfishly gave up his own research time to care for them as well as providing the emotional and financial support without which this project would have been impossible.

TABLE OF CONTENTS

	page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	vii
LIST OF FIGURES.	x
ABSTRACT	xii
Chapter	
I INTRODUCTION	1
The Purposes of the Study	3
Definitions	4
Outline of the Thesis	7
II REVIEW OF RESEARCH ON VISUAL SPATIAL ABILITIES .	9
Introduction	9
Spatial Subfactors.	10
Sex-Related Differences in Visual Spatial Abilities.	22
Another Type of Spatial Test.	28
An Alternate Framework.	34
Causes of Sex-Related Differences in Spatial Abilities.	41
III REVIEW OF RESEARCH ON MATHEMATICAL PROBLEM SOLVING ABILITY.	45
Introduction.	45
Sex-Related Differences in Mathematical Problem Solving Ability.	46
The Relationship Between Mathematical Problem Solving Ability and Visual Spatial Abilities.	65

Chapter		Page
IV	DESIGN OF THE STUDY	89
	Introduction	89
	Rationale for This Study	89
	Selection and Description of the Sample.	92
	Selection of the Tests of Visual Spatial Abilities	95
	Construction of the Test of Mathematical Problem Solving	103
	Scoring Procedures	108
V	CARRYING OUT THE STUDY.	110
	Introduction	110
	The Pilot Study and Resulting Changes.	110
	Gathering the Data for the Main Study.	117
	Scoring Procedures	118
VI	DATA ANALYSIS	131
	Introduction	131
	Sex-Related Differences in Performance on the Tests	132
	Relationships Between Visual Spatial Abilities and Mathematical Problem Solving Ability and Sex-Related Differences in These Relationships	150
VII	SUMMARY AND CONCLUSIONS	183
	Introduction	183
	Summary.	183
	Conclusions of the Study	184
	Limitations of the Study	193
	Suggestions for Future Research.	196
	Educational Implications	197
	Conclusion	198
	REFERENCES	200
	APPENDIX A: Sample Items from a Number of Figural Tests . .	215
	APPENDIX B: Tests Used in This Study.	226

	page
APPENDIX C: Descriptive Statistics from Other Administrations of the Tests Used in This Study	248
APPENDIX D: Scatterplots of Two Methods of Scoring the SR-O Tests.	252
APPENDIX E: Correlation Matrices.	253

Do Not File

LIST OF TABLES

Table		page
2.1	<u>t</u> Ratios for Sex Differences in Spatial Test Means in the Project Talent Study.	25
3.1	Sex-Related Differences in Applications Scales, National Longitudinal Study of Mathematical Abilities, X-, Y- and Z-Populations.	56
3.2	Sex-Related Differences by Analysis Scales, National Longitudinal Study of Mathematical Abilities, X-, Y- and Z-Populations.	56
3.3	Correlations Among Three Abilities in the CAA Project (1970): Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	67
3.4	Correlations Among Three Abilities in the CAA Project (1971): Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	69
3.5	Correlation Coefficients of Spatial Tests and Analysis Scales in the NLSMA	71
3.6	Per Cents of Total NLSMA Analysis and Application Scales on Which Significant Main Effects of Spatial Variables Were Observed.	74
3.7	Relationships Among Three Abilities in the Project Talent Study: Mathematical Problem Solving, Figural Reasoning, and Spatial Ability	77
3.8	Correlation Coefficients with Concealed Figures and Problem Solving in Students at Swarthmore and Yale	80
5.1	Descriptive Statistics from the Pilot Study Results.	112
5.2	Descriptive Statistics for Several Different Eighth-Grade Groups Taking the Spatial Tests	113
5.3	Textbook and Geometry Instruction for Each of the Mathematics Classes in the Sample.	119
5.4	Agreement over Coders 1, 2 and 3 on Solutions. . . .	125

Table		page
5.5	Agreement over Coders 1, 2, and 3 on Diagrams. . .	127
5.6	Agreement over Coders 1, 2, and 3 on Correct Representations.	129
6.1	Descriptive Statistics and Comparison of Means of Boys, Girls and Both for the Entire Test Battery	134
6.2	Reliability Estimates for Spatial Tests and Problem Scales for Boys, Girls and Both.	136
6.3	Problem Solving Item Analysis: Proportion of Correct Responses for Boys, Girls and Both	138
6.4	Distribution of Boys and Girls in Levels of Mathematics Classes.	140
6.5	<u>F</u> Values Generated by the Sex-by-Level ANOVAs. . .	142
6.6	Distribution of Boys and Girls in IQ Stanines. . .	145
6.7	<u>F</u> Values Generated by the Sex x IQ ANOVAs (Equal Cell Sizes).	147
6.8	Cell Means for the DAT Space Relations Test: Comparison of the Sex-by-IQ Subsample and the Whole Group.	148
6.9	Cell Means for the Cubes Comparison Test: Comparison of the Sex-by-IQ Subsample and the Whole Group.	149
6.10	The Relative Importance of Using Diagrams in Solving Three Types of Problems	154
6.11	Significant Sex-Related Differences in Correlation Coefficients	156
6.12	Correlations Between the Spatial Variables and Problem Solutions Total for the Whole Group. .	158
6.13	The Solutions Factor	161

Table		page
6.14	The Space Factor	163
6.15	The First Drawing Factor (Problem Types A and B) .	165
6.16	The Second Drawing Factor (Problem Type C)	167
6.17	Standardized Regression Coefficients	169
6.18	Standardized Regression Coefficients: Two- and Three-Dimensional Comparison	179
6.19	Canonical Correlation Analysis: Spatial Composite with Problem Solutions Composite	181

Do Not Film This Page

LIST OF FIGURES

Figure		page
1.1	Two Representations of a Rectangular Parallelepiped.	6
2.1	Guilford's Structure-of-Intellect Model	14
2.2	Figural Plane of the Structure-of-Intellect Model	14
2.3	A Translation in 2-space.	36
2.4	A Rotation in 2-space	36
2.5	A Reflection in 2-space	36
2.6	A Translation in 3-space	36
2.7	A Rotation in 3-space	36
2.8	A Reflection in 3-space	36
2.9	An Item Typical of the Cubes Test	39
4.1	A Matrix of Spatial Tests	96
4.2	A Set of Card Rotations Items	98
4.3	Two Cubes Comparison Items.	99
4.4	An Item from Form Board 2	100
4.5	An Item from DAT Space Relations.	102
4.6	A Matrix of Spatial Ability Tests Used in This Study	102
4.7	A Hidden Figures 2 Item	103
5.1	Example from the Instructions for the Cubes Comparison Test	115
5.2	An Example of a Figure Scored One for the Existence of a Diagram.	121

Figure		page
5.3	A Difficult Diagram to Score.	123
5.4	Some Solutions for Problem 6.	126
6.1	Sequence of Data Analysis, Part A, Sex-Related Differences in Performance.	133
6.2	Sex-by-Level Interaction: Cell Means on Diagram T ANOVA Using All Data.	143
6.3	Sequence of Data Analysis, Part B	152

THE INTERRELATIONSHIP OF SEX, VISUAL SPATIAL ABILITIES,
AND MATHEMATICAL PROBLEM SOLVING ABILITY IN GRADE SEVEN

Ann Koch Schonberger

Under the supervision of Professor Thomas A. Romberg

The intent of this study was to examine the hypothesis that males are more successful at solving mathematical and spatial problems than females. The relationship between visual spatial abilities and mathematical problem solving ability was also investigated. Specifically five questions were posed.

1. Are there sex-related differences in problem solving ability in mathematics?
2. Are there sex-related differences in any of the measures of visual spatial abilities?
3. What are the relationships between measures of visual spatial abilities and mathematical problem solving ability? Do these relationships differ by sex?
4. Are the relationships between visual spatial abilities and mathematical problem solving ability different for problems requiring differing amounts of spatial knowledge?
5. Is a subject's use of diagrams when solving mathematical problems related to successful solution of mathematical or spatial problems?

The sample consisted of the seventh grade students from one junior high school in Bangor, Maine. Each student took five spatial tests and a test of mathematical problems. The spatial tests were Card Rotations, Form Board 2, and Hidden Figures 2 from the NLSMA battery, Cubes Comparison from the ETS Kit, and Space Relations from the Differential Aptitude Test. The test of mathematical problems was constructed by the author from an item pool developed by Zalewski. Problems were chosen to make three subtests of problems differing in amount of spatial or geometric content. Level of mathematics class and IQ were also recorded for each subject. Questions about sex-related differences in performance were investigated using the t test; additional refinement was made possible by analysis of variance using the data on IQ and level of mathematics class. The relationships between solving spatial problems, drawing diagrams, and solving mathematical problems were explored using a variety of correlational techniques; these included factor analysis, multiple regression analysis and canonical correlation analysis.

Chapter VII

SUMMARY AND CONCLUSIONS

Introduction

In this chapter the plan and execution of the study and its results are summarized and interpreted. An effort is made to relate the results of this study to others in the field. Limitations of the study and possible remedies are also suggested. Finally, there is a discussion of the implications for educational practice and research.

Summary

The study was designed to search for sex-related differences in performance on tests of visual spatial abilities and mathematical problem solving. Questions were also asked about the relationships between those two types of abilities, and whether or not there are sex-related differences in the relationships. To investigate these questions, 176 seventh-grade students at the Fifth Street Junior High School in Bangor, Maine, were given a battery of five tests of visual spatial abilities and a test of mathematical problem solving. The spatial tests included both two- and three-dimensional tests of the Vz and SR-O subfactors as well as a test of visual dis-embedding. Problems for the mathematics test were chosen to make

three subtests of problems differing in the amount of spatial or geometric knowledge they involved. Data were also gathered on the level of mathematics class, teacher, textbook, and IQ for each student. The testing took place in late May and early June of 1975.

Of the 176 subjects tested there were complete test data on 170. This reduced sample was used in all the statistical analyses except the descriptive statistics on the tests. In order to test for sex-related differences in performance, t tests were performed on the boys' and girls' means for all the tests. In order to control for uneven sex distribution in levels of mathematics class and IQ stanines, sex-by-level and sex-by-IQ ANOVAs were carried out. The examination of the relationships between the spatial variables and mathematical problem solving and the search for sex-related differences in those relationships involved a number of statistical techniques. In addition to computing correlation matrices for both boys and girls, factor analysis, regression analysis, and canonical correlation analysis were employed. The results of all these statistical analyses are discussed in the next section in the context of previous and concurrent research.

Conclusions of the Study

The conclusions to be discussed here can be organized best in terms of the hypotheses stated in Chapter IV.

1. Boys and girls do not differ in their ability to solve mathematical problems.

The bulk of the literature on sex-related differences in the ability of seventh- and eighth-grade students to solve mathematical problems, reviewed in Chapter III, indicates that boys are better than girls at solving such problems. This seemed true especially at the upper-ability levels. The review also indicates that the differences may be limited to geometric problems or practical problems with spatial content.

The problem-solving study conducted by Meyer (1975) on fourth-grade students used a measure newly designed for her study by Romberg and Wearne (1975). Each item has three parts--one question testing understanding of the information in the problem, one testing application of the information in a routine way, and a third testing non-routine application or problem solving. By t tests, Meyer (1976) found no significant sex-related differences on any of three subscales or on the total. Fennema and Sherman (1976) used the Test of Academic Progress to measure mathematics achievement and found sex-related differences in only two of the four high schools participating in the study.

In this sample, none of the differences in the problem-solutions scales evaluated by the t test were significant at the .05 level. However, in all three sex-by-level ANOVAs the boys' performance on C-type problems was significantly better than the girls' performance at least at the .05 level. There was no sex-by-level interaction on any of the four solution scales, and examination of the cell means showed that the difference was not confined to the Level 1 (upper-

ability level) students. Since the one item which had shown significant sex bias was a C-type problem, the statistics were computed again after eliminating this item. The sex-related differences were reduced to nonsignificance. The results of this study and the concurrent research indicate that sex-related differences in problem solving ability found in earlier studies have disappeared in lower grades and are not as prevalent in high school populations as they once were.

2. There are no sex-related differences in performance on measures of any of the visual spatial abilities: SR-O or Vz, two- or three-dimensional, visual disembedding.

Some of the studies reviewed in Chapter II showed sex-related differences occurring in or before Grade 7. In the CAA Project's 1970 sample there was no sex-related difference on Spatial Relations, a form-board test, but in the 1971 sample the boys' performance on that test, as determined by t tests, was superior at the .01 level. Meyer (1976) gave a number of the cognitive abilities tests designed for the CAA Project and found that in her fourth-grade sample boys performed better than girls on Spatial Relations ($p < .01$). This was the largest sex-related difference found on any of the cognitive abilities tests; the only other difference reaching significance was in favor of boys on one other test with pictorial stimuli. Other studies of preadolescents summarized by Maccoby and Jacklin (1974) found few sex-related differences in tests of visual spatial abilities including visual disembedding.

Sex-related differences in visual spatial abilities become more apparent in high school students. There were small but significant differences on the spatial tests given by Project Talent (Flanagan, et al., 1964), and the DAT manual's norms show differences (again small) between boys and girls on its Space Relations test. However, in a recent study Fennema and Sherman (1976) found significant sex-related differences on DAT Space Relations in only two of the four high schools participating in the study.

The results of this study fit between the results of those cited above. Sex-related differences in Form Board 2 were significant at the .01 level just as they were in two of the three samples taking the form-board type test used by the CAA Project. Boys and girls performed equally well on DAT Space Relations, Hidden Figures 2, and Card Rotations. There is an indication that boys' performance on Cubes Comparison was better but the difference was not significant. With the exception of Form Board 2 there seems to be little difference at this age.

As on tests of mathematical problems it appears that the sex-related differences in performance on spatial tests are not as widespread as earlier studies indicated. In the recent studies the major differences observed were on form-board type tests. It may be that these tests (and perhaps also Cubes Comparison) represent a dividing point on the continuum of complexity hypothesized in Chapter II. If so, perhaps more boys than girls were able to visually transform the images which is a faster method of solving form-board items. Further research using form-board tests would be useful.

3. There are significant positive relationships between each of the visual spatial abilities and mathematical problem solving ability. The relationships are stronger for Vz tests than for SR-O tests and stronger for three-dimensional tests than for two-dimensional tests. These relationships do not differ by sex.

These hypotheses were investigated in a number of different ways. The correlation coefficients between all the spatial variables and the problem-solutions scales were significant and positive for girls and for the whole group. For boys many, but not all, of the correlation coefficients were significant. The problem-solutions scales of the mathematics tests had low but significant loadings for girls but not for boys on the Space factor in the factor analysis. In the regression analysis the spatial variables also predicted more of the variance on the solutions scales for girls than for boys. Hidden Figures 2 had high loadings on the Solutions factor for girls but not for boys. Likewise it contributed more to the regression equations for girls than for boys.

The fact that mathematical problem solving ability is more closely related to visual spatial abilities for girls than for boys was suggested by the results of the CAA Project. When Meyer (1976) factor analyzed her data separately by sex she found that Spatial Relations was relevant to two factors for girls but to only one for boys and for the whole group. This second factor, which appeared only for girls, seemed to involve mathematical problem solving: for example, the Number Series test was relevant to it and the third

problem solving scale also had significant loadings. In this study of seventh-grade students the relationship between the two abilities appeared to be closer for girls than for boys also, as was discussed in the preceding paragraph. These results all support the author's hypothesis about the sources of sex-related differences in spatial performance and factor structure. If girls as a group are less able to visualize the transformations involved in spatial tests and more often resort to reasoning about the figures, then the cognitive processes used in solving spatial items would be more like those used in solving mathematics problems for girls than for boys. Since problem-solving style was not investigated directly in this study, other explanations are possible. It should also be noted that few sex-related differences in the relationships between spatial and mathematical abilities have been found in studies with older subjects including the recent study of Fennema and Sherman (1976).

Another hypothesis that was generated in Chapter III by the results of the NLSMA and the Project Talent study is that Vz tests are better predictors of mathematical problem solving than SR-O tests. This appears to be true to some extent in this study. The DAT Space Relations test (Vz) often has the largest coefficient in the regression equations for the problem solution scales, especially for boys. On the other hand, Cubes Comparison, supposedly an SR-O test, usually makes a substantial contribution to the regression equations especially for girls. Given the fact that several different methods have been reported for solving Cubes items, it might

be hypothesized that Cubes Comparison is a Vz test for this sample. However, Form Board 2, supposedly also a Vz test, is the least valuable predictor in all of the regression equations except those involved with diagrams for C-type problems.

Another hypothesis is that DAT Space Relations and Cubes Comparison are better predictors because they varied more. This would explain the relative unimportance of Form Board 2 which had little variability, but not the importance of Hidden Figures 2 which had even less variance. A third possibility is that it is the three-dimensional nature of DAT Space Relations and Cubes Comparison that makes them good predictors of mathematical problem solving. This last possibility was supported by the results of the regression analyses.

The suggestion was made in the introductory chapter and again in the third chapter that poorer performance by girls on tests of mathematical problem solving might be due to deficiencies in spatial skills. This was only partially supported in the recent study by Fennema and Sherman (1976). In one of the schools participating in their study there were significant ($p < .01$) sex-related differences in favor of boys in both the DAT Space Relations test and their measure of mathematics achievement. In another school boys performed better than girls on DAT Space Relations at the .05 level but there were no significant differences on the mathematics measure. In a third school there were sex-related differences in favor of boys ($p < .01$) on the mathematics measure but none on the spatial test.

In both schools in which the boys did better on the mathematics measure, there were significant sex-related differences on over half of the eight affective measures; this was not true in the other two schools. The results of this study also argue against attributing sex-related differences in mathematical problem solving to differences in visual spatial abilities. The one test of this present study on which there were significant sex-related differences in favor of boys, Form Board 2, was the least important in the regression equations for both sexes.

4. Each type of visual spatial ability is more closely related to solving mathematical problems with high spatial content than to solving problems with little spatial content.

The curious result of this part of the investigation was that the solution of B-type problems was more closely related to the spatial variables than the solution scale for A-type problems especially for girls. This appeared in the factor analysis where Problem Solutions B loaded substantially on the Space factor and in the regression analysis where the spatial variables predicted more of the variance on Problem Solutions B than on Problem Solutions A or C. One possible explanation for this occurrence involves the nature of the stimuli for the A- and B-type problems. Five of the eight A-type problems have pictorial stimuli; by definition none of the B-type problems do. It may be that in solving B-type problems subjects used spatial skills more often to visualize the real situation in the problem or to organize the given information

than they did in A-type problems where the visual organization was already presented. However, this hypothesis is purely speculative.

5. Use of a diagram in solving a mathematical problem is positively related both to the ability to solve that problem and to visual spatial abilities. There are no sex-related differences in these relationships.

The factor analysis showed that the relationship between drawing a diagram for a problem, especially a diagram that represents the information correctly, and solving that problem is closer for A-type problems than for the other types. Comparison of correlation coefficients for solution of A-, B-, and C-type problems with their respective drawing scales also supports this conclusion. This was expected because of the way in which the problems were categorized. What was unexpected was that Level 1 girls would score higher on the diagrams and correct representations scales than Level 1 boys while at Level 2 there was no sex-related difference. One noticeable difference between the Level 1 and Level 2 classes was the amount of geometry taught; only one Level 2 class had had any geometry whereas all the Level 1 classes had been exposed to some geometry that year. It may be that girls, who are supposedly more successful at school, were applying what they had learned about geometry more often and more successfully. The fact that this difference is significant for the A-type problems but only a trend for the other types supports this explanation.

The regression analysis enables one to examine the relationship between the spatial variables and the problem scales involved with diagrams. What is surprising is that the spatial variables are better predictors of the solution scales than of the three types of scales related to diagrams except for Correct Representations A. However, the spatial variables are better predictors of correctly representing a problem with a diagram than they are of simply drawing a diagram. This seems again to indicate that these spatial tests involve a substantial reasoning component especially for this seventh-grade sample.

Limitations of the Study

The conclusions drawn from any piece of educational research should always be tempered by the limitations of the study. The discussion of the limitations of this study is divided into three parts: limitations due to the sample, limitations in the measures used, and limitations of the design. Post hoc suggestions for improving the study will be made in each section.

The lack of random selection of subjects places a strain on the assumptions necessary for statistical analysis and the implications of the results for other groups. The sample was described in some detail in Chapter IV and the problem of inference left to the reader. A random sample from a larger population would have been preferable.

Both the tests of visual spatial abilities and the test of mathematical problem solving had their limitations. The means on

Form Board 2 and Hidden Figures 2 were lower in the main study than in the pilot study. Especially on Hidden Figures 2 the variability was seriously restricted by the difficulty of the test; over 90 subjects had scores of zero, one, or two. Both Hidden Figures 2 and Form Board 2 might have been better predictors of mathematical problem solving ability if they had not been so difficult for these subjects. In a future study of seventh-grade students the easier forms of these two tests from the NLSMA battery might be used.

Although the reliability in Cubes Comparison was increased in the main study, the experimenter was still concerned about the possible difficulty with the instructions. For students whose logical skills are not well-developed, differentiating between "could be the same" and "must be different" might be difficult. There is a similar test by Werdelin (see Appendix A) in which the task is to decide which letters are on opposite faces of a cube given three drawings of the cube. It was not used in this study because no descriptive statistics were available for that test for such young subjects. However, in a future study it might be pilot-tested for use.

The test of mathematical problem solving yielded the same low means in the main study as it did in the pilot test and in Zalewski's (1974) samples. This is not necessarily bad since problem solving is a complex behavior. However, there were a few indications that the subjects did not understand the information in some problems,

especially the geometry problems. A test like the Romberg-Wearne test used by Meyer (1975) on which each item tested understanding of the information in the problem and its routine application before posing the "real" problem would be preferable. Scoring would also be facilitated by using a multiple choice format. With the set of plausible distractors generated by this administration of the test, such a measure might not be too difficult to construct. As was discussed in Chapter VI, Item 3 should be eliminated because of its sex bias and its ambiguity.

The limitation of the design was that common to most correlational studies. Although information was gathered on the existence and strength of relationships between problem solving in mathematics and visual spatial abilities, the cognitive processes used in solving mathematical and spatial problems were not determined. The results of the study, as well as the review of research, suggest that the essence of the relationship is some sort of figural reasoning ability. Meyer's (1975) results support this position. A future study of junior high school students using a similar design might include a figure matrix test in order to see if figural reasoning ability is the common ground between mathematical problem solving ability and spatial skills.

Suggestions for Future Research

Ultimately, however, it seems that the relationship between visual spatial abilities, figural reasoning ability, and mathematical problem solving ability will become known only through subjects' own accounts of their methods of solving items on those types of tests. Barrett (1953) and French (1965) have used this method to study the way in which college students solve spatial items. Kilpatrick (1967), Lucas (1972), and Zalewski (1974) have developed and used process-sequence coding systems to record the results of students' thinking aloud while solving mathematical problems. Newell and Simon (1972) have designed computer programs to simulate human problem solving in mathematics as well as in other areas. A study in which subjects' reports of their own problem-solving styles on a variety of spatial items and mathematical problems were coded by a process-sequence system would be very enlightening. In particular such a research strategy might answer the question of whether or not sex-related differences in performance on spatial tests are related to different methods of solving spatial items.

The lack of sex-related differences in the ability to solve mathematical problems in this study and in Meyer's may be confined to these samples. Certainly replications are necessary. However, the results of this study show the importance of checking the tests used for sex-biased items. The recent study by Fennema and Sherman (1976) as well as the older study by Carey (1955) indicated that sex-related differences in performance on mathematics tests coincide with

differences in attitudes towards mathematics and problem solving. A second study by Fennema and Sherman (in progress) using middle-school subjects as well as a study by Whitaker (in progress) should provide additional information on this relationship.

Educational Implications

The fact that recent studies have shown few, if any, sex-related differences in mathematical problem solving ability should caution educators against retaining the myth that boys excel at mathematical reasoning. Such myths affect teachers' expectations for boys and girls, and students may try to live up to these expectations. Such sex-biased views might be partly responsible for the smaller number of women electing advanced mathematics courses in high school or college. The results of this study also show the importance of constructing tests which are free of sex bias. The test used in this study was designed to be neutral with respect to male and female actors in the problems. The results of the item analysis indicate that problem content should also be attended to in constructing new tests of mathematical problems.

Spatial tests are used by schools almost exclusively for career guidance. Although a few sex-related differences in spatial tests appeared in recent studies, the differences were not as frequent as one would expect from a review of earlier research. Certainly such differences are not great enough to recommend different academic programs or careers for boys and girls as groups. If it is true

that spatial skills are important in such prestigious careers as architecture and engineering, then spatial training ought to be part of the school curriculum. More spatial training in schools would probably benefit both boys and girls. This is a relatively untouched area of curriculum research and development.

Although sex-related differences in visual spatial abilities did not seem to cause sex differences in mathematical problem solving ability, spatial skills were related to the problem solving process. Although some mathematics programs for elementary schools emphasize pictorial as well as symbolic representation of mathematical problems, this is not common. It should be encouraged, especially in light of the low means found in this study and in other problem solving research (Meyer, 1975; Zalewski, 1974). The Russian literature (e.g., Botsmanova, 1960/1972) is suggestive of ways in which children can learn to use diagrams more effectively to solve mathematical problems. Schools in the United States might experiment with these visual methods.

Conclusion

This investigation of the ability of seventh-grade students to solve mathematical and spatial problems found almost no difference between boys' and girls' capabilities. It appears that in 1975, officially designated International Women's Year, the inferior position which has been attributed to females in those intellectual areas should be modified, just as the position of women has been challenged in legal, social and economic matters. This is not to

say that equality has been achieved. However, the changes that have occurred suggest that further change is possible. With sensitivity and dedication mathematics educators should be able to do their part in order that women may participate fully in our mathematical society.

REFERENCES

- Ahmann, J. S. Consumer math: selected results from the first national assessment of mathematics (NAEP Mathematics Report No. 04-MA-02). U. S. Government Printing Office, 1975.
- Anglin, L., Meyer, R., & Wheeler, J. The relationship between spatial ability and mathematics achievement at the fourth- and sixth-grade levels. Working Paper No. 115. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1975.
- Barrett, E. S. An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. Journal of Psychology, 1953, 36, 17-25.
- Bennett, G. K., Seashore, H. G., & Wesman, A. G. Differential aptitude tests: Administrator's handbook. New York: The Psychological Corporation, 1973.
- Berry, J. W. Temne and Eskimo perceptual skills. Journal of International Psychology, 1966, 1, 207-299.
- Berry, P. C. An exploration of the interrelations among nonintellectual predictors of achievement in problem solving. Technical Report No. 4. New Haven, Connecticut: Department of Psychology and Department of Industrial Administration, Yale University, 1958.

- Berry, P. C. A second exploration of the interrelations among nonintellectual predictors of achievement in problem solving. Technical Report #5. New Haven, Connecticut: Department of Psychology and Department of Industrial Administration, Yale University, 1959.
- Bieri, J., Bradburn, W. M. & Galinsky, M. D. Sex differences in perceptual behavior. Journal of Personality, 1958, 26, 1-12.
- Bock, R. D. & Kolakowski, D. Further evidence of sex-linked major-gene influence on human spatial visualizing ability. American Journal of Human Genetics, 1973, 25 (1), 1-14.
- Boe, B. L. A study of the ability of secondary pupils to perceive plane sections of selected solid figures. Mathematics Teacher, 1968, 61, 415-421.
- Botsmanova, M. E. [The forms of pictorial visual aids in instruction in arithmetic problem solving] (J. Kilpatrick & I. Wirszup, Eds., J. W. Teller, trans.). Soviet studies in the psychology of learning and teaching mathematics (Vol. VI). Stanford, California: School Mathematics Study Group, Stanford University & Survey of Recent East European Mathematical Literature, University of Chicago, 1972. (Originally published, 1960.) (a)
- Botsmanova, M. E. [The role of graphic analysis in solving arithmetic problems] (J. Kilpatrick & I. Wirszup, Eds. & J. W. Teller, trans.). Soviet studies in the psychology of learning and teaching mathematics (Vol. VI). Stanford, California: School Mathematics Study Group, Stanford University & Survey of Recent East European Mathematical Literature, University of Chicago, 1972. (Originally published, 1960.) (b)

Bruner, J. S. Some theorems on instruction illustrated with reference to mathematics. In E. R. Hilgard (Ed.), Theories of learning and instruction, 63rd Yearbook (Part 1) of the National Society for the Study of Education.

Burt, C. The structure of the mind: a review of the results of factor analysis. British Journal of Educational Psychology, 1949, 19, 176-199.

Carey, G. L. Reduction of sex differences in problem solving by improvement of attitude through group discussion. Technical Report #9. Stanford, California: Stanford University Department of Psychology, 1955.

Carey, G. L. Sex differences in problem solving-performance as a function of attitude differences. Journal of Abnormal and Social Psychology, 1958, 56, 256-260.

Carnegie Commission on Higher Education. Opportunities for women in higher education. New York: McGraw-Hill Book Company, 1973.

Carry, L. R. Interaction of visualization and general reasoning abilities with instructional treatment in algebra (Doctoral dissertation, Stanford University, 1968). Dissertation Abstracts International, 1968, 29A, 475. (University Microfilm Order #68-11,280.)

Carry, L. R. Patterns of mathematics achievement in grades 7 and 8: X-population. In J. W. Wilson, L. S. Cahen & E. G. Begle, (Eds.), NLSMA Report #11. Stanford, California: School Mathematics Study Group, Stanford University Press, 1970.

- Carry, L. R. & Weaver, J. F. Patterns of mathematics achievement in grades 4, 5 and 6: X-population. In J. W. Wilson, L. S. Cahen, & E. G. Begle (Eds.), NLSMA Report #10. Stanford, California: School Mathematics Study Group, Stanford University Press, 1969.
- Chetverukhin, N. F. [An experimental investigation of pupils' spatial concepts and spatial imagination] (J. Kilpatrick & I. Wirszup, Eds. & J. W. Teller, trans.). Soviet studies in the psychology of learning and teaching mathematics (Vol. V). Stanford, California: School Mathematics Study Group, Stanford University & Survey of Recent East European Mathematical Literature, University of Chicago, 1971. (Originally published, 1949.)
- Clark, R. B., Coladarci, A. P. & Caffrey, J. Statistical Reasoning and Procedures. Columbus, Ohio: Charles E. Merrill Books, Inc., 1956.
- Dodson, J. W. Characteristics of successful insightful problem solvers. In J. W. Wilson & E. G. Begle (Eds.), NLSMA Report #31. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972.
- Downie, N. M. & Heath, R. W. Basic Statistical Methods (3rd ed.). New York: Harper & Row, Publishers, 1970.
- Droege, R. C. Sex differences in aptitude maturation during high school. Journal of Counseling Psychology, 1967, 14, 407-411.
- Eastman, P. M. & Carry, L. R. Interaction of spatial visualization and general reasoning abilities with instructional treatment in quadratic inequalities. Journal for Research in Mathematics Education, 1975, 6, 142-149.

- Fehr, H. F., The secondary school mathematics improvement study: a unified mathematics program. The Mathematics Teacher, 1974, 67 (1), 25-33.
- Fennema, E. Mathematics learning and the sexes: a review. Journal for Research in Mathematics Education, 1974, 5 (3), 126-139. (a)
- Fennema, E. Mathematics, spatial ability and the sexes. Paper presented at the meeting of the American Educational Research Association, Chicago, 1974. (b)
- Fennema, E., & Sherman, J. Sex-related differences in mathematics learning: Myths, realities and related factors. Paper presented at the meeting of the American Association for the Advancement of Science, Boston, February, 1976.
- Flanagan, J., et. al. The American High School Student Today. Pittsburgh: University of Pittsburgh, 1964.
- French, J. W. The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monograph No. 5, 1951.
- French, J. W. The relationship of problem-solving styles to the factor composition of tests. Educational and Psychological Measurement, 1965, 25, 9-28.
- French, J. W., Ekstrom, R. B. & Price, L. A. Kit of Reference Tests for Cognitive Factors (Revised 1963). Princeton, New Jersey: Educational Testing Service, 1969.
- Fruchter, B. Measurement of spatial abilities: history and background. Educational and Psychological Measurement, 1954, 14, 387-395.
- Garai, J. E. & Scheinfeld, A. Sex differences in mental and behavioral traits. Genetic Psychology Monographs, 1968, 77, 169-299.

- Glennon, V. J. & Callahan, L. G. A guide to current research: elementary school mathematics. Washington, D. C.: Association for Supervision and Curriculum Development, 1968.
- Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill Book Company, 1967.
- Guilford, J. P., Fruchter, B. & Zimmerman, W. S. Factor analysis of the army air forces, Sheppard Field battery of experimental aptitude tests. Psychometrika, 1952, 16, 45-68.
- Guilford, J. P. & Zimmerman, W. S. The Guilford-Zimmerman Aptitude Survey. Beverly Hills, California: Sheridan Psychological Service, 1948.
- Hadamard, J. The psychology of invention in the mathematical field. New York: Dover, 1954.
- Harris, M. L. & Harris, C. W. A structure of concept attainment abilities. Wisconsin Monograph Series. Madison: Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, 1973.
- Hayes, W. L. Statistics for the social sciences (2nd ed.). New York: Holt, Rinehart and Winston, Inc., 1973.
- Heimer, R. T. & Lottes, J. J. Toward a theory of sequencing: an integrated program of research. Journal for Research in Mathematics Education, 1973, 4 (2), 85-93.
- Hilton, T. L. & Berglund, G. W. Sex differences in mathematics achievement--a longitudinal study. The Journal of Educational Research, 1974, 67, 231-237.
- Hobson, J. R. Sex differences in primary mental abilities. Journal of Educational Research, 1947, 41, 126-132.

- Hoffman, L. R. & Maier, N. R. F. Social factors influencing problem solving in women. Journal of Personality and Social Psychology, 1966, 4, 382-390.
- Horner, M. Toward an understanding of achievement related conflicts in women. In Stacey, J., Bereaud, S., & Daniels, J. (Eds.), And Jill Came Tumbling After: Sexism in American Education. New York: Dell Publishing Company, Inc., 1974.
- Jackson, D. N., Messick, S. & Meyers, C. T. Evaluation of group and individual forms of embedded figures for field-independence. Educational and Psychological Measurement, 1964, 24, 177-192.
- Jarvis, O. T. Boy-girl differences in elementary school arithmetic. School Science and Mathematics, 1964, 64, 657-659.
- Kabanova-Meller, E. N. [The role of the diagram in the application of geometric theorems] (J. Kilpatrick & I. Wirszup, Eds. & M. Ackerman, trans.). Soviet studies in the psychology of learning and teaching mathematics (Vol. IV). Stanford, California: School Mathematics Study Group, Stanford University & Survey of Recent East European Mathematical Literature, University of Chicago, 1970. (Originally published, 1950.)
- Kagan, J. & Kogan, N. Individual variation in cognitive processes. In P. H. Mussen (Ed.), Oarmichael's manual of child psychology (3rd ed.) Vol. 1. New York: Wiley, 1970.

Kilpatrick, J. Analyzing the solution of word problems in mathematics: an exploratory study (Doctoral dissertation, Stanford University, 1967). (University Microfilm No. 68-6442)

Kilpatrick, J. & McLeod, G. K. Patterns of mathematics achievement in grade 9: Y-population. In J. W. Wilson, L. S. Cahen & E. G. Begle (Eds.), NLSMA Report #13. Stanford, California: School Mathematics Study Group, Stanford University Press, 1971. (a)

Kilpatrick, J. & McLeod, G. K. Patterns of mathematics achievement in grade 11: Y-population. In J. W. Wilson, L. S. Cahen & E. G. Begle (Eds.) NLSMA Report #15. Stanford, California: School Mathematics Study Group, Stanford University Press, 1971. (b)

Kleinfeld, J. S. Intellectual strengths in culturally different groups: an Eskimo illustration. Review of Educational Research, 1973, 43 (3), 341-359.

Kline, W. E. A synthesis of two factor analyses of intermediate algebra. Psychometrika, 24 (3), 343-359.

Lucas, J. F. An exploratory study in the diagnostic teaching of heuristic problem solving strategies in calculus (Doctoral dissertation, University of Wisconsin, 1972).

Maccoby, E. E. Sex differences in intellectual functioning. In E. E. Maccoby (Ed.), The development of sex differences. Stanford, California: Stanford University Press, 1966.

Maccoby, E. E. & Jacklin, C. N. The psychology of sex differences. Stanford, California: Stanford University Press, 1974.

- McLeod, G. K. & Kilpatrick, J. Patterns of mathematics achievement in grades 7 and 8: Y-population. In J. W. Wilson, L. S. Cahen & E. G. Begle, NLSMA Reports #12. Stanford, California: School Mathematics Study Group, Stanford University Press, 1969.
- McLeod, G. K. & Kilpatrick, J. Patterns of mathematics achievement in grade 10: Y-population. In J. W. Wilson, L. S. Cahen & E. G. Begle (Eds.), NLSMA Report #14. Stanford, California: School Mathematics Study Group, Stanford University Press, 1971.
- Menchinskaya, N. A. [Intellectual activity in solving arithmetic problems] (J. Kilpatrick & I. Wirszup, Eds. & D. A. Henderson, trans.). Soviet studies in the psychology of learning and teaching psychology (Vol. III). Stanford, California: School Mathematics Study Group, & Survey of Recent East European Mathematical Literature, University of Chicago, 1969. (Originally published, 1946.)
- Meyer, R. A. A study of the relationship of mathematical problem solving performance and intellectual abilities of fourth-grade children. Doctoral dissertation, University of Wisconsin-Madison, 1975.
- Meyer, R. A. A study of the relationship of mathematical problem solving performance and intellectual abilities of fourth-grade boys and girls. Working Paper No. 160. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1976.

- Michael, W. B., Guilford, J. P., Fruchter, B. & Zimmerman, W. S. The description of spatial-visualization abilities. Educational and Psychological Measurement, 1957, 17, 185-199.
- Milton, G. A. The effects of sex-role identification upon problem solving skill. Journal of Abnormal and Social Psychology, 1957, 55, 219-244.
- Milton, G. A. Five studies of the relation between sex role identification and achievement in problem solving. Technical Report #3. New Haven, Connecticut: Department of Psychology and Department of Industrial Administration, Yale University, 1958.
- Mitchelmore, M. C. Cross-cultural research on concepts of space and geometry. Paper presented at the Research Workshop on Space and Geometry, Georgia, May, 1975.
- Parsley, K. M., Powell, M. & O'Connor, H. A. Further investigation of sex differences in achievement of under-, average and over-achieving students within five IQ groups in grades four through eight. Journal of Educational Research, 1964, 57, 268-270.
- Parsley, K. M., Powell, M., O'Connor, H. A. & Deutsch, M. Are there really sex differences in achievement? Journal of Educational Research, 1963, 56, 210-212.
- Poincaré, H. [Intuition and logic in mathematics] (G. B. Halstead, trans.). Mathematics Teacher, 1969, 62, 205-212. (Originally published in English in 1929.)
- Polya, G. Mathematical discovery: on understanding, learning and teaching problem solving (Vol. 1). New York: Wiley, 1962.

- Romberg, T. A. & Wilson, J. W. The development of the tests. In J. W. Wilson & E. G. Begle (Eds.), NLSMA Report #7. Stanford, California: School Mathematics Study Group, Stanford University Press, 1969.
- Romberg, T. A. & Wilson, J. W. Patterns of mathematics achievement in grade 12: Z-population. In J. W. Wilson & E. G. Begle (Eds.), NLSMA Report #18. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972.
- Sells, L. W. High school math as a critical filter in the job market. (ERIC Document Reproduction Service No. ED 080 351).
- Sheehan, T. J. Patterns of sex differences in learning mathematical problem solving. Journal of Experimental Education, 1968, 36 (summer), 84-87.
- Sherman, J. A. Field articulation, sex, spatial visualization, dependency, practice, laterality of the brain and birth order. Perceptual and Motor Skills, 1974, 38, 1223-1235.
- Sherman, J. A. Problem of sex differences in space perception and aspects of intellectual functioning. Psychological Review, 1967, 4, 290-299.
- Sherrill, J. M. The effects of different presentations of mathematical word problems upon achievement of tenth-grade students. School Science and Mathematics. 1973, 73, 277-282.
- Simon, H. A. & Newell, A. Human problem solving. American Psychologist, 1971, 26, 145-159.
- Smith, I. M. Spatial ability. London: University of London Press Ltd., 1964.

- Stafford, R. E. Sex-differences in spatial visualization as evidence of sex-linked inheritance. Perceptual and Motor Skills, 1961, 13, 428.
- Stanley, J. C., Keating, D. P. & Fox, L. H. Mathematical Talent: Discovery, Description and Development. Baltimore: The Johns Hopkins University Press, 1974.
- Sweeney, E. J. Sex Differences in Problem Solving. Technical Report #1. Stanford, California: Stanford University Department of Psychology, 1953.
- Thurstone, L. L. A factorial study of perception. Chicago: University of Chicago Press, 1944.
- Thurstone, L. L. Some primary abilities in visual thinking. Chicago: Psychometric Laboratory, Report No. 59, University of Chicago, 1950.
- Thurstone, T. G. Manual for the SRA primary mental abilities tests--ages 11 to 17. Cleveland: Science Research Associates, 1958.
- Tyler, L. E. The psychology of human differences. New York: Appleton-Century-Crofts, 1965.
- Very, P. S. Differential factor structures in mathematical ability. Genetic Psychology Monographs, 1967, 75, 169-207.
- Wearne, D. Development of a test of mathematical problem solving which yields a comprehension, application, and problem solving score. Doctoral dissertation, University of Wisconsin-Madison, in progress.

- Webb, L. F. & Carry, L. R. Interaction of spatial visualization and general reasoning abilities with instructional treatment in quadratic inequalities: a follow-up study. Journal for Research in Mathematics Education, 1975, 6, 132-140.
- Webb, L. F. & Sherrill, J. M. The effects of differing presentations of mathematical word problems upon the achievement of preservice elementary teachers. School Science and Mathematics, 1974, 74, 559-565.
- Werdelin, I. The mathematical ability. Lund, Sweden. C. W. K. Gleerup, 1958.
- Werdelin, I. Geometrical ability and the space factor in boys and girls. Lund, Sweden: C. W. K. Gleerup, 1961.
- Whitaker, D. A study of the relationships between selected non-cognitive factors and the problem solving performance of fourth-grade children. Doctoral dissertation, University of Wisconsin-Madison, in progress.
- Willoughby, R. H. Field-dependence and locus of control. Perceptual and Motor Skills, 1967, 24, 671-672.
- Wilson, J. W. Patterns of mathematics achievement in grade 10: Z-population. In J. W. Wilson & E. G. Begle (Eds.), NLSMA Report #16. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972. (a)
- Wilson, J. W. Patterns of achievement in grade 11: Z-population. In J. W. Wilson, L. S. Cahen & E. G. Begle (Eds.), NLSMA Report #17. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972. (b)

- Wilson, J. W. & Begle, E. G. (Eds.) Correlates of mathematics achievement: summary. NLSMA Report #26. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972. (a)
- Wilson, J. W. & Begle, E. G. (Eds.) Intercorrelations of mathematical and psychological variables. NLSMA Report #33. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972. (b)
- Wilson, J. W., Cahen, L. S. & Begle, E. G. (Eds.). Descriptive and statistical properties of the X-population. NLSMA Report #4. Stanford, California: School Mathematics Study Group, Stanford University Press, 1968. (a)
- Wilson, J. W., Cahen, L. S. & Begle, E. G. (Eds.). Description and statistical properties of the Y-population. NLSMA Report #5. Stanford, California: School Mathematics Study Group, Stanford University Press, 1968. (b)
- Wilson, J. W., Cahen, L. S. & Begle, E. G. (Eds.). Description and statistical properties of the Z-population. NLSMA Report #6. Stanford, California: School Mathematics Study Group, Stanford University Press, 1968. (c)
- Wilson, J. W., Cahen, L. S. & Begle, E. G. (Eds.) Statistical procedures and computer programs. NLSMA Report #8. Stanford, California: School Mathematics Study Group, Stanford University Press, 1972.
- Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R. & Karp, S. A. Psychological Differentiation. New York: Wiley, 1962.

Wolfe, L. R. The effects of spatial visualization on spatial ability and arithmetic achievement in junior high school students.

(Doctoral dissertation, SUNY Albany, 1970). Dissertation Abstracts International, 1970, 31, 2801A. (University Microfilms No. 70-25, 486.)

Yakimanskaya, I. S. [Individual differences in solving geometry problems on proof] (J. Kilpatrick & I. Wirszup, Eds. & J. W. Teller, trans.). Soviet studies in the psychology of learning and teaching mathematics (Vol. IV). Stanford, California: School Mathematics Study Group, Stanford University & Survey of Recent East European Mathematical Literature, University of Chicago, 1970. (Originally published, 1959.)

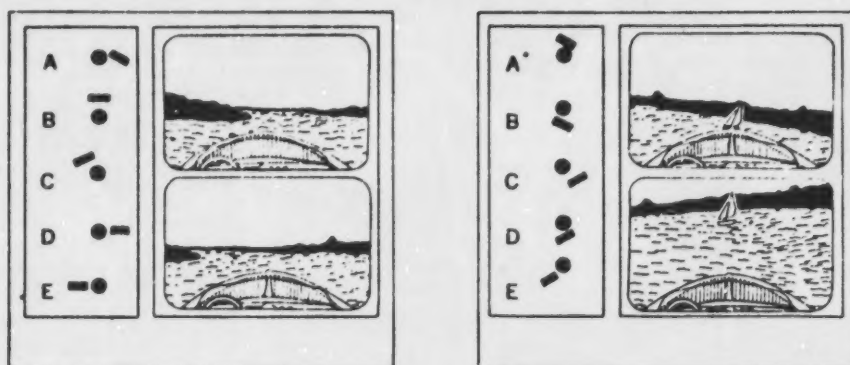
Zalewski, D. L. An exploratory study to compare two performance measures: an interview-coding scheme of mathematical problem solving and a written test (2 vols.). Technical Report No. 306. Madison: Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1974.

Zimmerman, W. S. Hypotheses concerning the nature of the spatial factors. Educational and Psychological Measurement, 1954, 14, 396-400.

Appendix A

SAMPLE ITEMS FROM A NUMBER OF FIGURAL TESTS

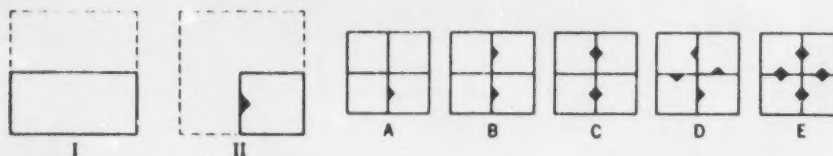
GUILFORD'S TESTS

A Spatial Orientation Test

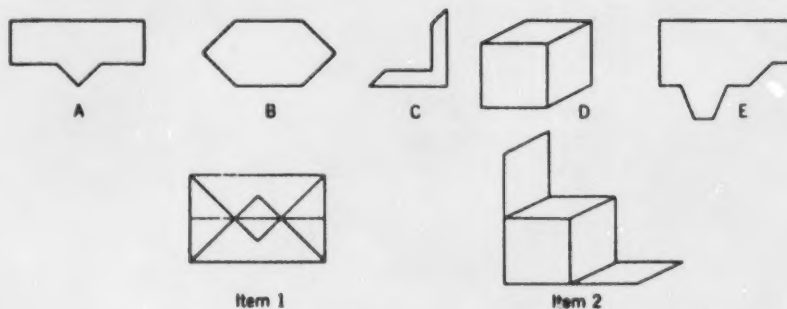
Two sample items from a spatial-orientation test. As the position of the boat changes with respect to the background when we go from the upper to the lower picture in each item, which alternative symbol should show what the change is like? The dot in each alternative answer indicates the position of the prow of the boat in the upper picture, and the rectangle indicates the position of the boat in the lower picture. *Answers: D and B, for first and second items, respectively. (From Part V of the Guilford-Zimmerman Aptitude Survey, courtesy of the Sheridan Psychological Services, Inc.)*

GUILFORD'S TESTS, continued.A Test of Visualization

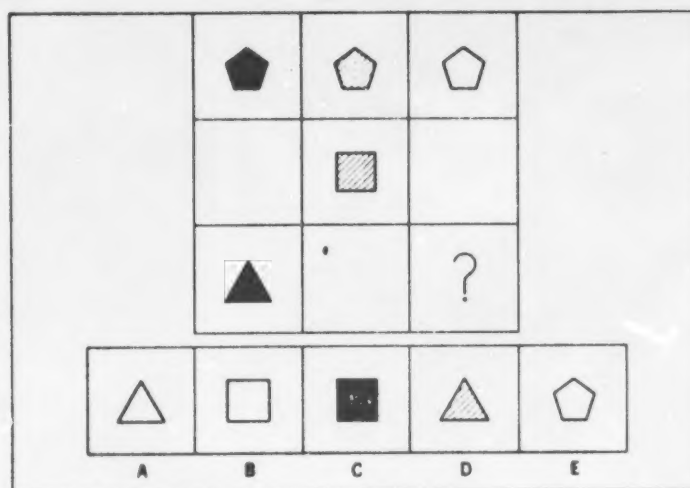
Two items from a test of visualization. In each item, imagine how the clock at the left is turned in accordance with the arrow or arrows shown on the sphere. Which clock at the right shows how the clock would look after making the turn or turns? The first and second turns in Item II must be made in the 1-2 order. *Answers: B and C, respectively. (Courtesy of the Sheridan Psychological Services, Inc., from Part VI of the Guilford-Zimmerman Aptitude Survey, Form B.)*

Spatial Visualization I

A sample item from Spatial Visualization I, an experimental AAF test based on a test designed by the writer (Guilford & Lacey, 1947). Diagrams I and II show two steps in folding a square piece of paper and cutting a notch in a certain location. Which alternative shows how the paper would look when unfolded? *Answer: C.*

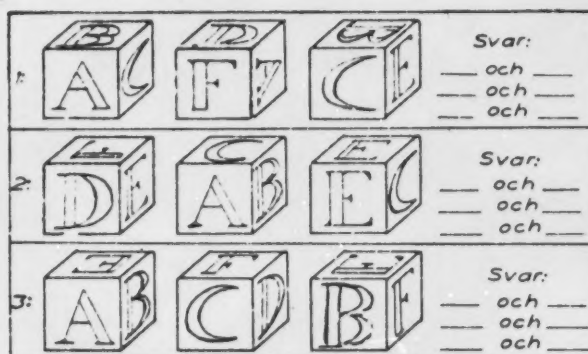
GUILFORD'S TESTS, continued.Hidden Figures

Sample items from the Hidden Figures test. Which of the five simple figures at the top is concealed in each of the item figures?
 Answers: 1, A; 2, D. (This test is based upon figures originally designed by Gottschaldt, 1926.)

Figure Matrix

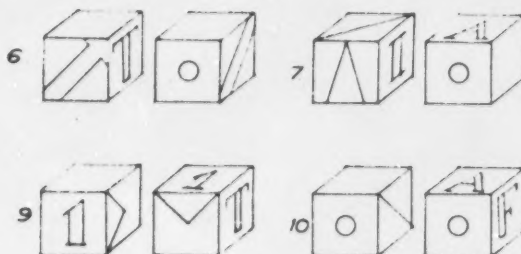
A sample item from the Figure Matrix test. What kind of figure should appear in the cell with the question mark? Answer: A.

WERDELIN'S TESTS

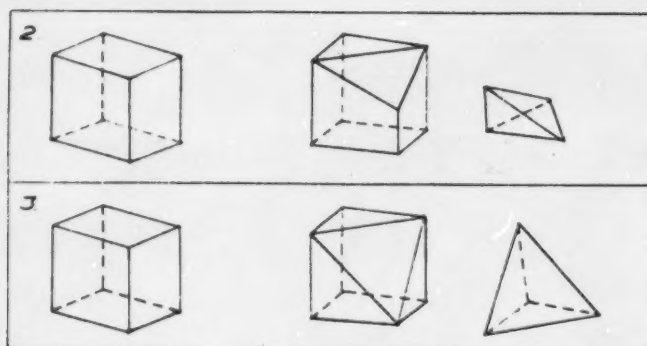
Cubes

This test contains one example and 16 tasks. In each item there are three drawings of cubes, and the subjects are told that they represent the same cube. On the faces of the cubes are letters, different letters on each of the faces. The subjects are asked to determine which letters are on parallel (opposite) faces.

Scoring: 2 points if three pairs of faces are correct, 1 point if two pairs of faces are correct and one wrong, or one pair of faces correct and none wrong.

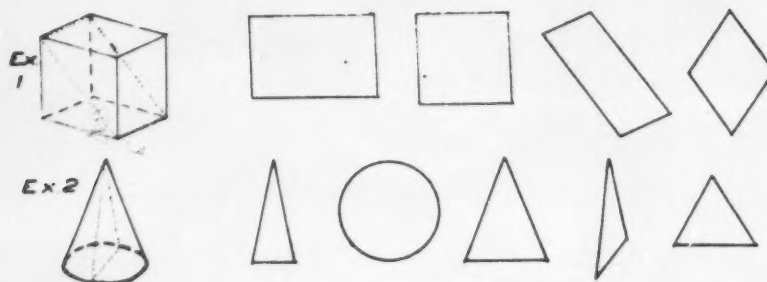
Pairs of Cubes

There are 42 pairs of cubes, including 2 examples. The subjects are instructed to compare drawings of cubes to see whether they can represent the same cube or must necessarily represent two different cubes. They are told that there is a different design on each face of the cube, and that the cubes may be seen from different positions.

WERDELIN'S TESTS, continued.Solid Figures

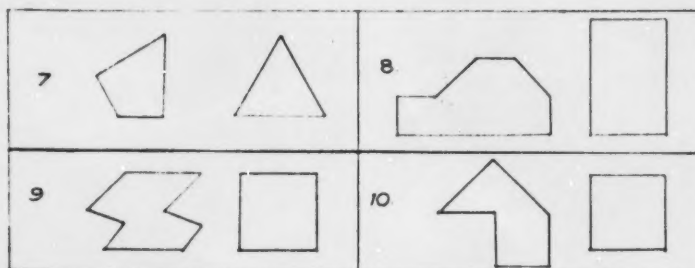
This test consists of one example and 12 tasks. The following oral instructions were given: "On the left in each item there is a cube or polygon, on the right two smaller solids. You are to show, by drawing lines on the figure, how to divide it by means of a plane to get the two figures on the right."

Scoring: number of correct answers. Time: 6-1/2 minutes.

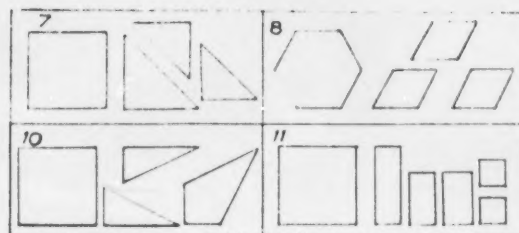
Sections

To the left in each item is a picture of a solid, and to the right are five surfaces. The subjects were told that each solid is intersected by a plane. The lines where the plane intersects the faces of the solid are indicated by dotted lines. Then the subjects were told to imagine how the new surface would look if one looked at it perpendicularly to the surface. They were told to mark one of the five surfaces.

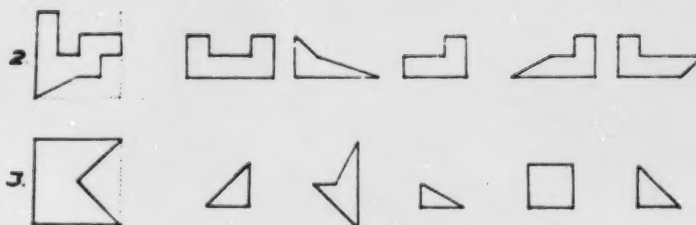
Scoring: number of correct answers. The test is untimed.

WERDELIN'S TESTS, continued.Form Board I

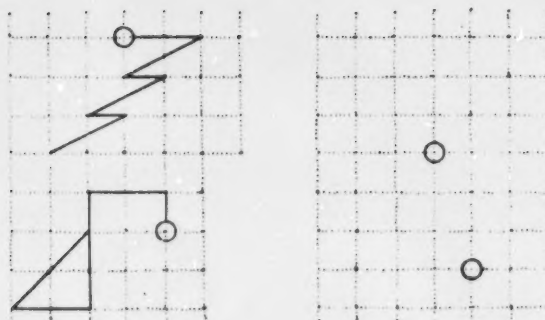
This test consists of 2 examples and 14 actual items. The score is the number of correct items, and the subjects are given 6 minutes for the test. The following oral instructions are given: "On the left in each square is an irregular figure, on the right is a more regular figure: square, triangle, or rectangle. You are to divide the irregular figure into two parts, by means of one straight line, so that the parts can be put together to form the regular figure on the right. You are to draw the straight line."

Form Board II

The task on this test is to divide the figure on the left in order to get the pieces shown on the right.

WERDELIN'S TESTS, continued.Puzzles

The subjects are instructed to indicate which two figures on the right complete the square at the left in each row. The score is the number of correct rows. Time: 6 minutes.

Inversion

The subjects are told to draw the standard figure on the left upside down, beginning at the small ring. The test is scored by the number of items exactly correct. Time: 6-1/2 minutes.

WERDELIN'S TESTS, continued.Figures

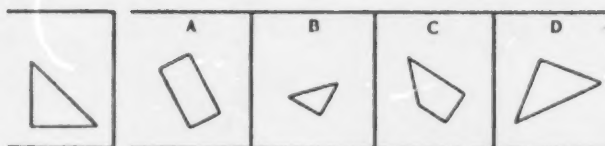
⁶ Σ	Z Ẑ ∩ ∨ N S Z
⁷ E	∇ ∪ ∩ ∇ ∇ ∇ ∇
⁸ 4	∪ ∩ ∩ ∩ ∩ ∩ ∩

The test consists of 24 rows of "meaning-less" geometrical figures. Three of these rows are used as examples. The figure on the left in each row is a standard figure, with which the other seven figures are to be compared. Some of these figures are exactly the same shape as the standard figure, but turned round; some of them are the same shape except that they are turned round and reflected. The subjects are instructed to mark the figures which were reflected in comparison with the standard figure. The test is scored by the number of correct rows. Time: 5 minutes.

CAA PROJECT TESTS

Spatial Relations

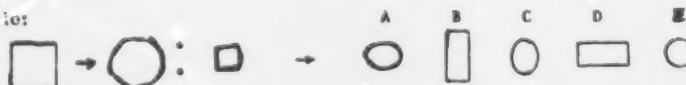
Example:



This is the Spatial Relations test from the PMA 4-6 test battery. It was reproduced for use by the CAA Project with permission of the publisher, Science Research Associates, Inc. From four choices the subject chooses the figure that would complete a given figure to form a square.

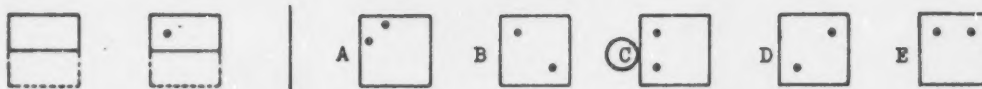
Figure Analogies

Example:



This is an analogies type test using figures rather than words. The test consists of 22 selected items from the 33 items of the Non-Verbal Test No. 3 of the Lorge-Thorndike Intelligence Tests (1964). These items were reproduced with the permission of the publisher, Houghton Mifflin Company. Items which contained pictures of things instead of figures and items which were of a series type, such as an increase in size of each succeeding figure, were not used.

ETS KIT TEST

Paper Folding

Part I of this test was used by the NLSMA and titled Punched Holes. The two figures at the left show how a piece of paper has been folded and punched. The task is to decide which of the five figures at the right shows where the holes will be when the paper is completely unfolded. Each part had 10 items and the time for each was 3 minutes.

Appendix B

TESTS USED IN THIS STUDY

CARD ROTATIONSINSTRUCTIONS

The following figure represents a triangle-shaped card.



All of the following figures represent the same card. The card has been turned around on the same page to different positions.



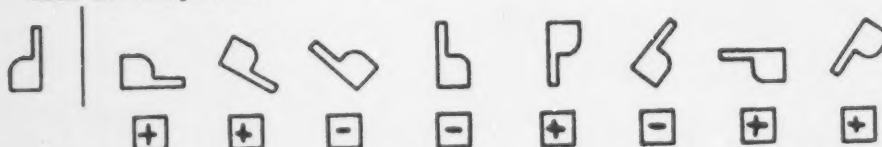
Now look at the card below.



This, too, is the same card. However, it has been turned over. This is the only way that it could be put in this position. It could not be put in this position by just turning it around. In the same way all of the following figures represent the same card after it has been turned over.



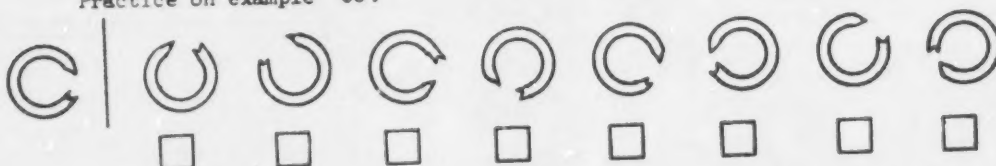
Look at example 0.



The figure to the left of the line represents a card. The eight on the right represent the same card. In some of these, the card has just been turned

around and in others it has been turned over. You are to put a plus (+) in the box below the figure if it has been turned around and a minus (-) in the box if it has been turned over. Example 0 has been marked for you.

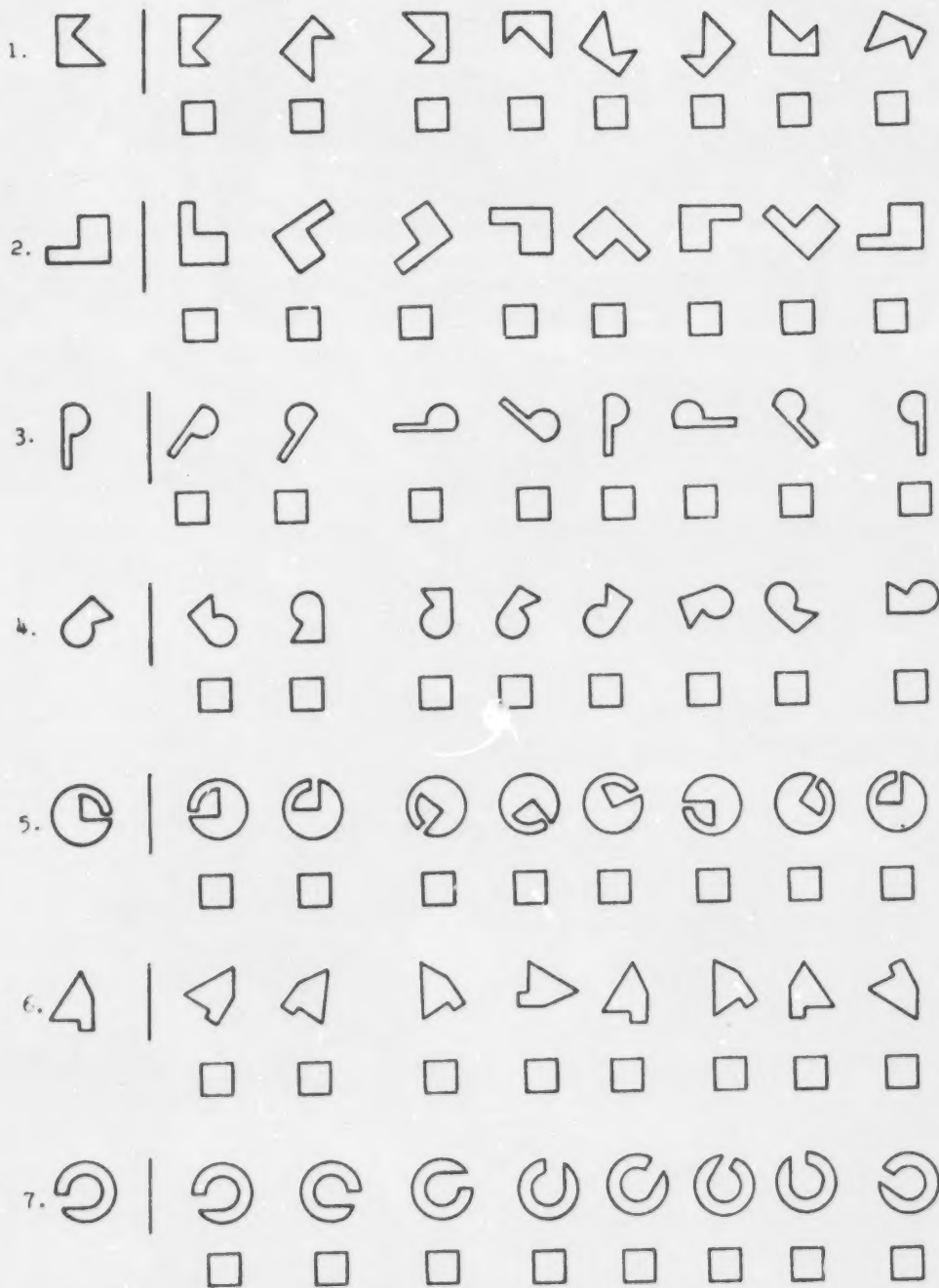
Practice on example 00.



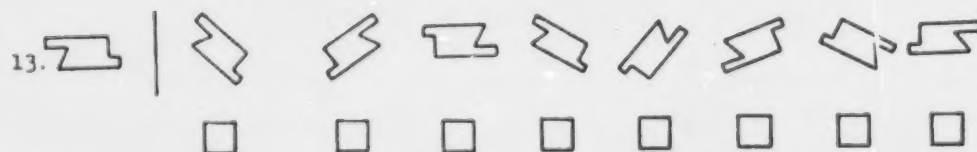
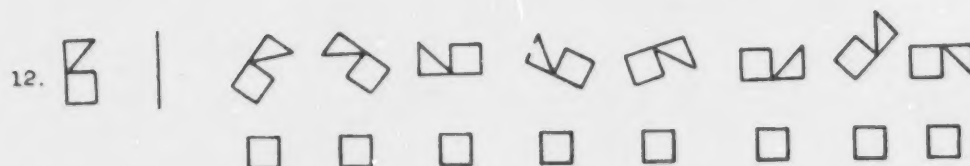
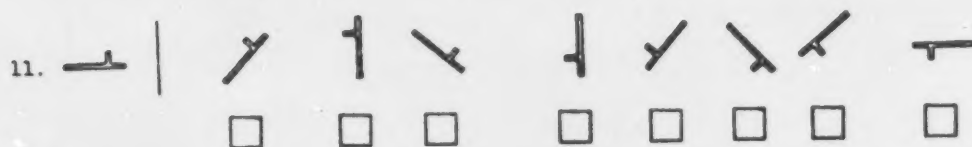
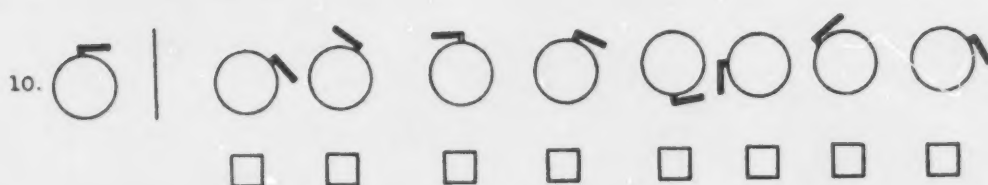
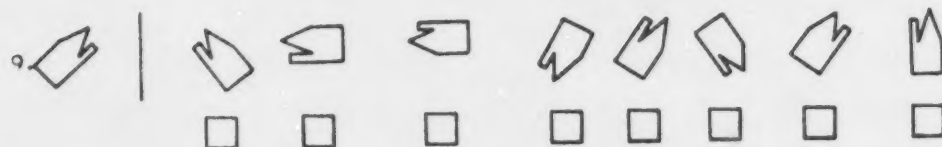
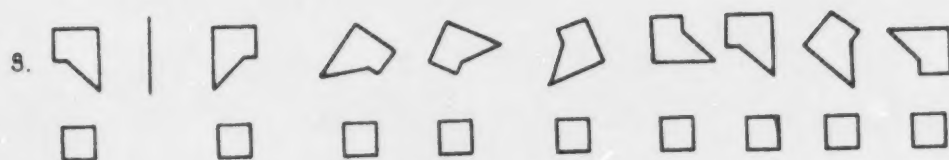
Work as quickly and as accurately as you can. Do not guess wildly.

You will have 4 minutes for this section. There are two pages in this section. When you have finished, STOP. Please do not go on until you are asked to do so.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.



GO ON TO THE NEXT PAGE.

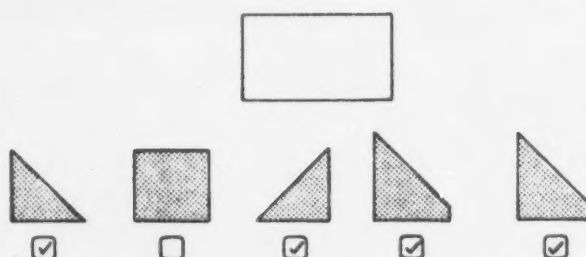


STOP STOP. If you finish before time is called, check your work on this part. Do not go back to earlier sections.

FORM BOARD
INSTRUCTIONS

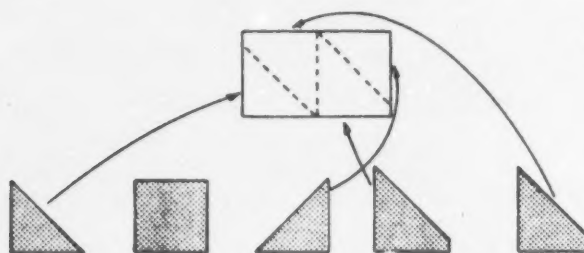
In this section you are to tell which pieces can be put together to make a certain figure.

Look at Example 0.



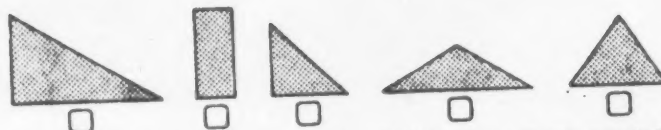
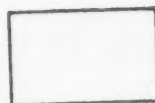
You are to decide which of the five shaded pieces can be put together to form the rectangle.

One way to do the problem is to form the rectangle as follows:

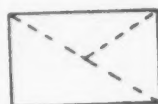


When you know which pieces make the complete figure check the boxes under pieces that are used. See how it has been done in Example 0.

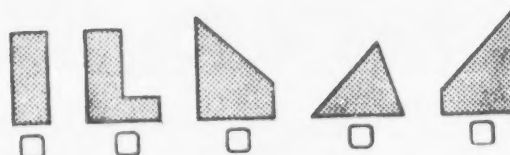
Now try Example 00.



In Example 00, the first, fourth, fifth pieces are needed. You should have checked the lines under these pieces. See how they fit together.



Now try Example 000:



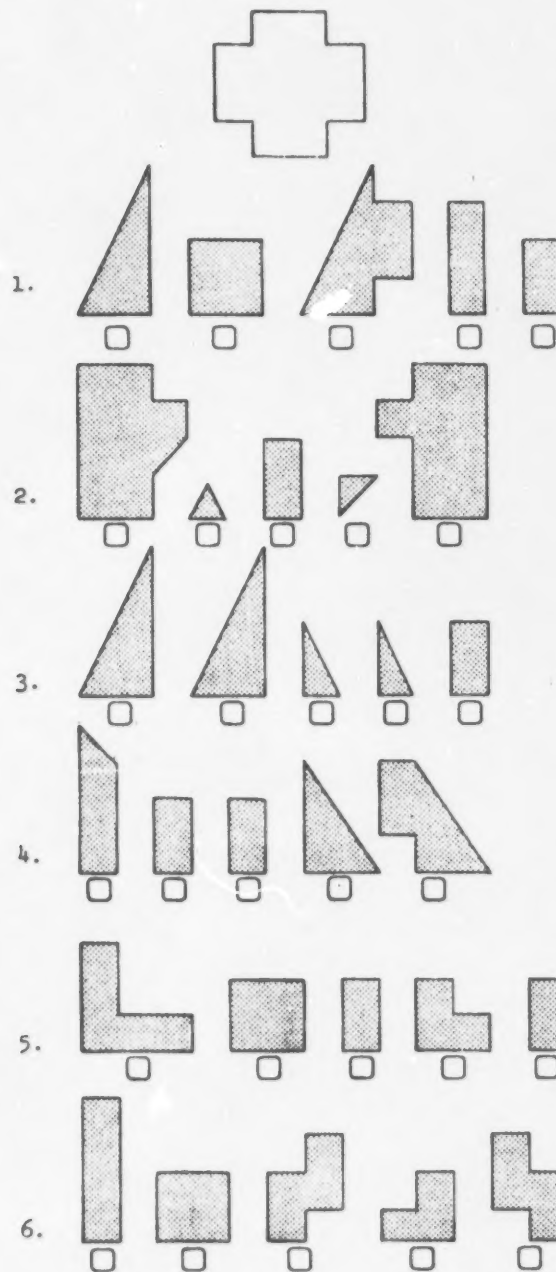
At the top of each page is a geometrical figure. Beneath each figure are several questions. Each question consists of a row of five shaded pieces. You are to decide which of the five shaded pieces will make the complete figure when put together. The number of shaded pieces needed to make the complete figure may be from two to five. Each piece may be turned around to any position but it cannot be turned over. It may help you to sketch the way the pieces fit together. You may use any blank space for doing this.

Work as quickly and as accurately as you can. Do not guess wildly.

You will have 8 minutes for this section. There are four pages in this section. When you have finished, STOP. Please do not go on until you are asked to do so.

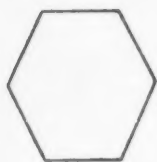
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

USE THIS SPACE
FOR SCRATCHWORK



GO ON TO THE NEXT PAGE

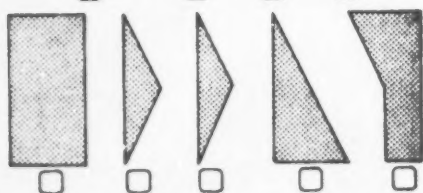
USE THIS SPACE
FOR SCRATCHWORK



7.



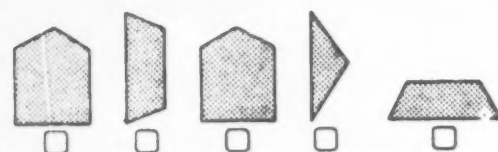
8.



9.



10.



11.



12.



GO ON TO THE NEXT PAGE

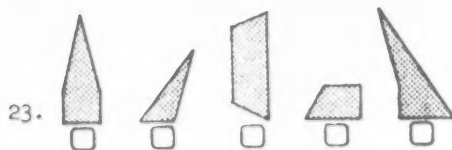
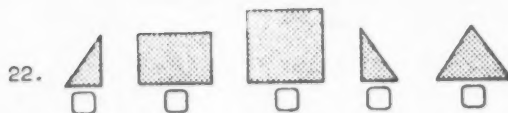
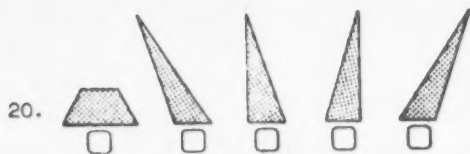
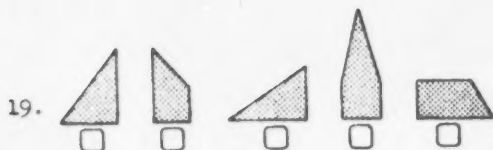


USE THIS SPACE
FOR SCRATCHWORK

13. ☐
14. ☐
15. ☐
16. ☐
17. ☐
18. ☐

GO ON TO THE NEXT PAGE

USE THIS SPACE
FOR SCRATCHWORK



STOP. If you finish before time is called, check your work on this part. Do not go back to earlier sections.

HIDDEN FIGURES

237

INSTRUCTIONS

In this test you have a pattern on the left. On the right there are five figures. You have to find which one of these five figures can be found in the pattern on the left.

Look at the sample question below. Circle the correct letter.



A



B



C



D



E

The correct answer is A as figure A is the only figure hidden in the pattern. The figure below shows how figure A is hidden in the pattern.



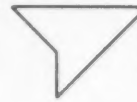
You will notice that there are some extra lines passing through this figure. These extra lines are to make the figures harder to find.

There will be only one figure in each pattern. It will always be the same size and shape. It will not be turned around or turned over.

Here is a sample question for you to try.



I



M



N



P

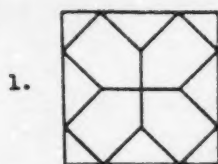


Q

Work as quickly and as accurately as you can. You should only guess if you can rule out some of the choices. Do not guess wildly.

You will have 10 minutes. There are three pages in this section. When you have finished the third page, STOP. Please do not go on until you are asked to do so.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.



A



B



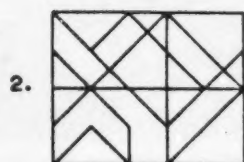
C



D



E



L



M



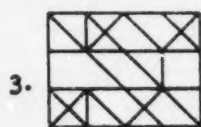
N



P



Q



A



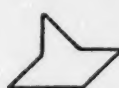
B



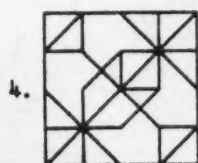
C



D



E



L



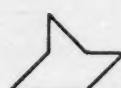
M



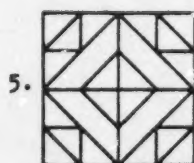
N



P



Q



A



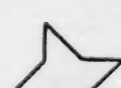
B



C

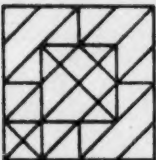
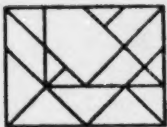
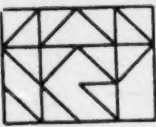
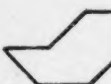
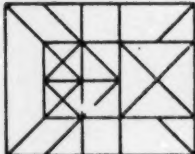
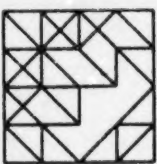
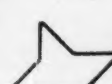
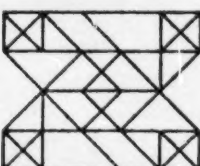


D

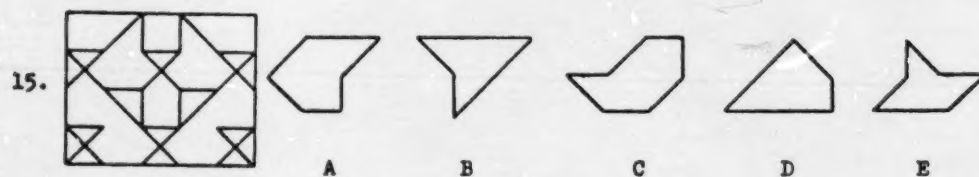
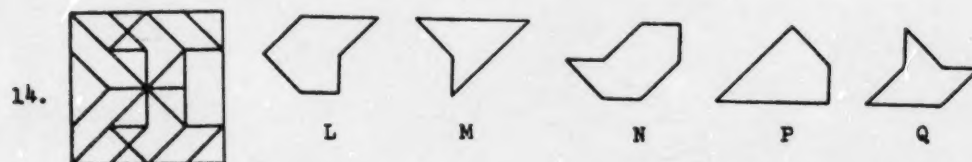
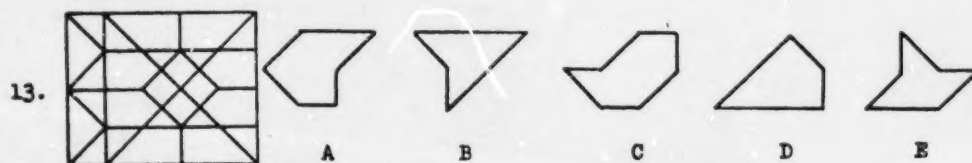
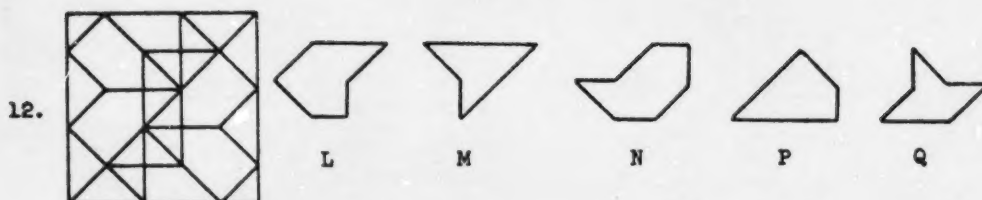


E

GO ON TO THE NEXT PAGE

6.      
L M N P Q
7.      
A B C D E
8.      
L M N P Q
9.      
A B C D E
10.      
L M N P Q
11.      
A B C D E

GO ON TO THE NEXT PAGE



If you finish before time is called, check your work on this part. Do not go back to earlier sections.

INSTRUCTIONS

Wooden blocks such as children play with are often cubical in shape and have a different letter or number on each of their six faces (top, bottom, and four sides). Each problem on the test consists of drawings of two cubes of this kind. Look at the two pairs of cubes below:

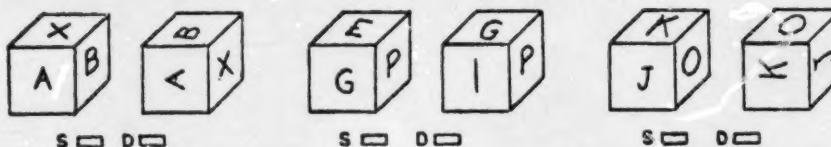


The first pair of cubes is marked D, because the two drawings must be of different cubes. If the left cube is turned so that the A is in front, the N on this cube will be hidden and will be on the side opposite from where it is shown on the right-hand cube.

The second pair is marked S, because the two drawings can represent the same cube. When the left cube is rolled over to the left, the X will no longer show, a C will appear, and the cubes can be the same.

RULE: When a face is hidden on one cube and is turned up on the second, the letter or number on that face is correct unless that letter or number has already been shown in a different position on either of the cubes.

Be sure you see that this pair can represent the same cube.



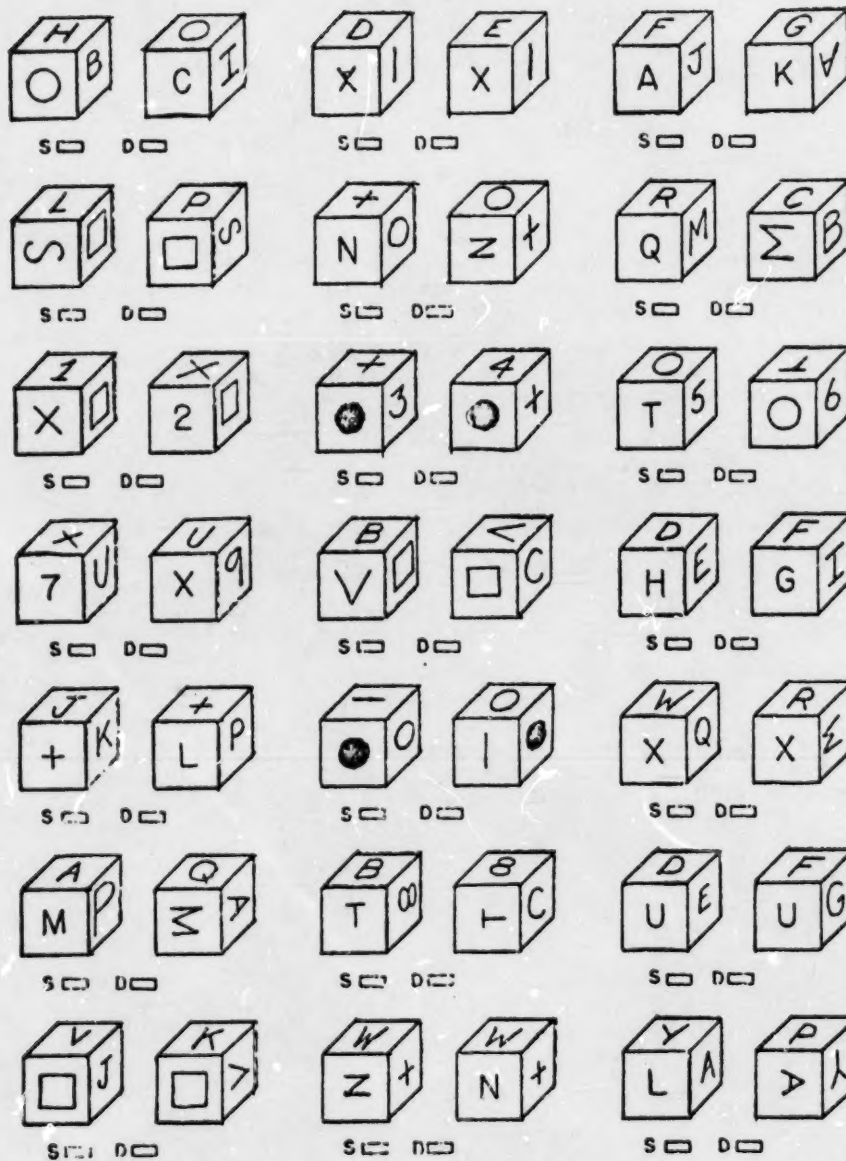
The first pair should be marked "different," because X cannot be next to the top of A on the left cube and the bottom of A on the right cube. The second pair is "different," because P has its side next to G on the left cube and its top next to O on the right cube. The third pair is "same."

Your score on this test will be the number marked correctly minus the number marked incorrectly. Therefore, it will not be to your advantage to guess unless you have some idea which choice is correct. Work as quickly as you can without sacrificing accuracy.

You will have 3 minutes for each of the two parts of this test. Each part has 1 page. When you have finished part 1, STOP.

DO NOT TURN THE PAGE UNTIL YOU ARE ASKED TO DO SO.

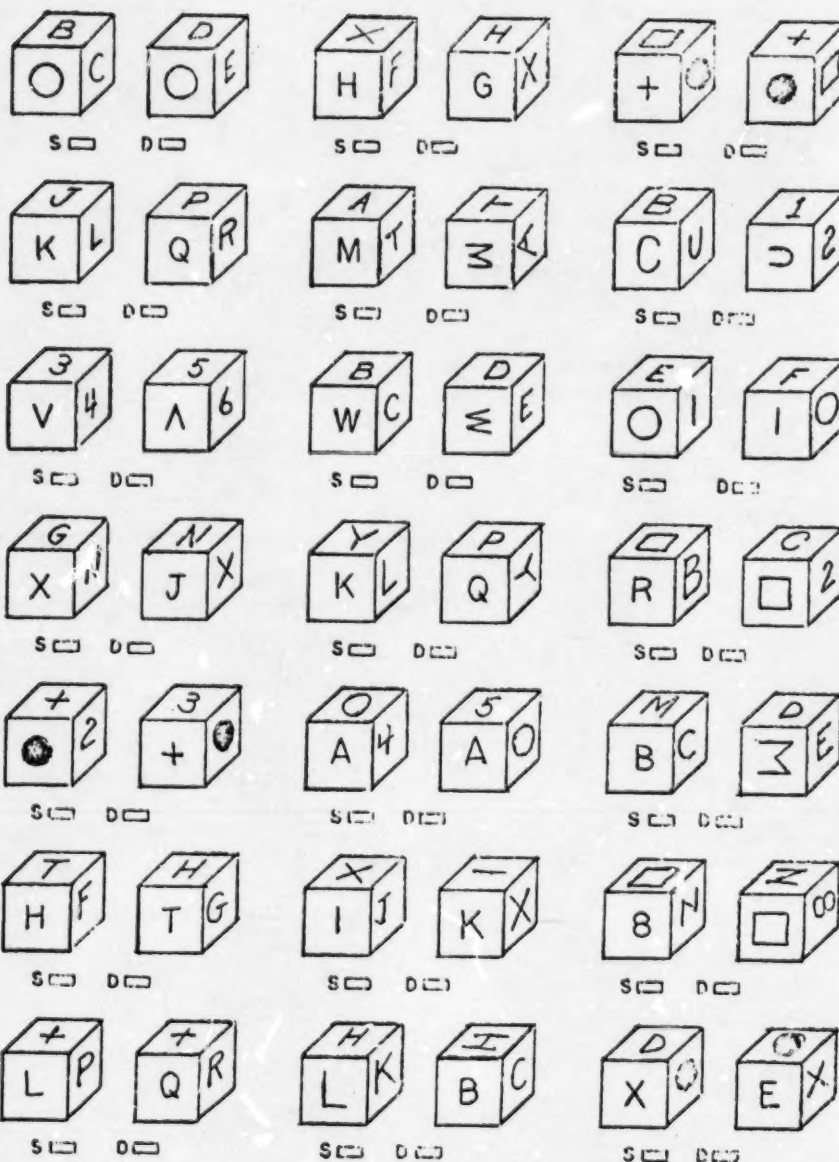
Part 1 (3 minutes)



DO NOT GO ON TO THE NEXT PAGE UNTIL ASKED TO DO SO.

STOP.

Part 2 (3 minutes)



DO NOT GO BACK TO PART 1 AND
DO NOT GO ON TO ANY OTHER TEST UNTIL ASKED TO DO SO.

STOP.

Mathematical Problems Tests

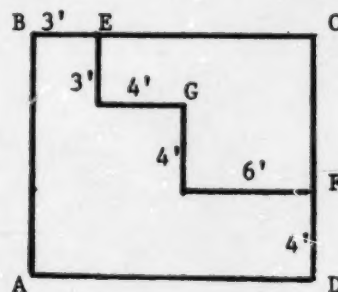
Problems Used in the Main Study

1. There was half of an apple pie left. Then Nate ate one half of the half and Kate ate one half of what Nate left. What part of the pie was left after Kate ate?
2. Two lines can cross at only one point, but three lines can cross at three points. What is the most points at which five lines can cross?

3. Which player has the best record when you consider both shots attempted and shots made?

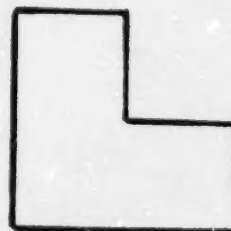
Name	Attempted	Made
Ray	12	7
Ann	8	5
Luke	10	5
Rose	8	4
Eli	15	7

4. What is the perimeter of (distance around) the rectangle ABCD?



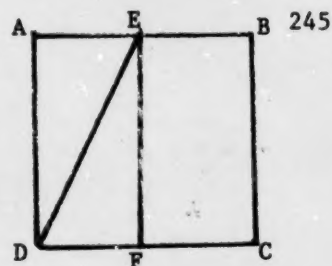
5. Four students are standing in the lunch line. In how many different ways (or orders) could these four students be lined up single file?

6. A mother and father owned three connected squares of land and wanted to divide it among their 4 children. Draw lines to show how they could divide up the land so each child gets an equal share.



7. The perimeter of (distance around) a swimming pool in the shape of a rectangle is 148 feet. If the length of the pool is 50 feet, how many square feet of surface does the pool have?
8. When Vincent answered 60 questions correctly on a test, he had $\frac{4}{5}$ of the answers right. How many questions were on the test?

9. ABCD is a square with E halfway between A and B, and F halfway between D and C. If each side of the square is ten inches, what is the area of triangle DEF?



10. If six bushels of wheat will plant four acres, how many bushels of wheat are needed to plant 30 acres?
11. John buys pencils at three for ten cents and sells them at a nickel each. How much profit would he earn on a dozen pencils?
12. What number comes next in 1, 2, 4, 7, 11, _____?
13. Tiles for floors come in different shapes. Which one of the shapes pictured here could not cover (without leaving spaces) a square floor?



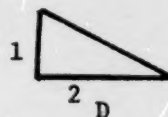
A



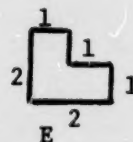
B



C



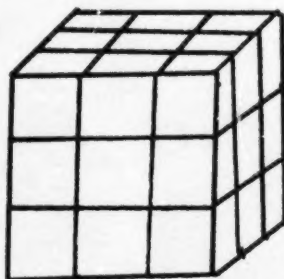
D



E

14. Jim left Bangor at 3:20 and took one hour and fifty minutes to drive to Portland. What time did he arrive in Portland?
15. A class of thirty students was divided into two groups. One group had eight more students than the other. How many students were in the larger group?
16. One taxi driver gets 35¢ for each dollar clocked on the taxi meter. She also gets tips. If she made a total of \$25.00 one day for clocking \$60.00 on the meter, how much money in tips did she get?
17. There were 18 brown-eyed students on the bus and 12 students had brown hair. If there was a total of 26 students on the bus, what is the smallest possible number of students that had both brown hair and brown eyes?
18. If 76 cookies fill five boxes with six cookies left over, how many of the same sized boxes will 100 cookies fill?
19. Six students belong to the basket ball team but only five can play at a time. How many different groups of five players can be formed by the six students?
20. What whole number must m be in order for $\frac{100}{m}$ to be a fractional number between 12 and 13?

21. Laura got scores of 63, 72, and 65 on her first three tests. What score must she get on her fourth test in order to end up with an average of 70 for the four tests?
22. Cindy borrowed \$3000 to buy a car. She agreed to pay \$100.00 a month for 3 years to repay the loan plus interest. How much interest did she pay in the 3 years?
23. The perimeter of (distance around) a rectangular flower garden is 60 feet. There is a $2\frac{1}{2}$ foot wide sidewalk around the garden. What is the perimeter of the outer edge of the sidewalk?
24. The large cube was painted red on all sides and then cut up into 27 smaller cubes. How many of the smaller cubes have exactly two red sides?



Problems in Category A: 2, 4, 6, 7, 9, 13, 23, 24

Problems in Category B: 1, 5, 8, 10, 15, 17, 18, 19

Problems in Category C: 3, 11, 12, 14, 16, 20, 21, 22

Appendix C
DESCRIPTIVE STATISTICS FROM OTHER ADMINISTRATIONS
OF THE TESTS USED IN THIS STUDY

Table 1
Card Rotations (NLSMA)

Sample	Sample Size	Mean	S. D.
X-Pop., Fall, Grade 5	422	41.85	18.35
Y-Pop., Fall, Grade 8	1565	54.98	20.83
Y-Pop., Fall, Grade 10	771	73.41	22.35
Y-Pop., Spring, Grade 10	741	85.55	21.40
Z-Pop., Fall, Grade 11	782	67.17	23.13

(Wilson, Cahen & Begle, 1968a, 1968b, 1968c)

Table 2
Form Board 2 (NLSMA)

Sample	Sample Size	Mean	S. D.	α	S. E.
Y-Pop., Fall, Grade 8	565	6.65	3.65	.76	1.77
Z-Pop., Fall, Grade 11	289	8.48	3.48	.75	1.74

Reliability coefficients (Cronbach's α) computed are coefficients of internal consistency. Since this test had a possible speed factor, the interpretation of both α and S. E. is questionable (Wilson, Cahen & Begle, 1968b, 1968c).

Table 3
Hidden Figures 2 (NLSMA)

Sample	Sample Size	Mean	S.D.	α	S. E.
X-Pop., Fall, Grade 8	1091	4.37	2.58	.60	1.64
Y-Pop., Fall, Grade 8	1589	4.01	2.48	.59	1.58
Y-Pop., Fall, Grade 11	694	6.64	3.58	.80	1.62
Z-Pop., Fall, Grade 11	826	5.11	2.84	.69	1.59

There is a possible speed factor in this test so the Cronbach's α and S. E. should be interpreted with caution (Wilson, Cahen & Begle, 1968a, 1968b, 1968c).

Table 4
DAT Space Relations

Sample	Sample Size	Mean	S. D.	Reliability
Boys, Fall, Grade 8	7000 +	24.1	9.9	.94
Girls, Fall, Grade 8	6900 +	23.0	9.0	.95

(Bennett, Seashore & Wesman, 1973)

Table 5
Cubes Comparison

Sample	Mean	S.D.
Vietnamese Helicopter Trainees (page 1 only)	7.3	6.50
31 US Army Trainees	16.4	8.5
College Age Women (page 1)	5.63	4.64
(page 2)	5.24	5.57
156 Eighth Graders (both pages)	9.81	2.96
46 College Undergraduates (both pages)	22.74	9.37

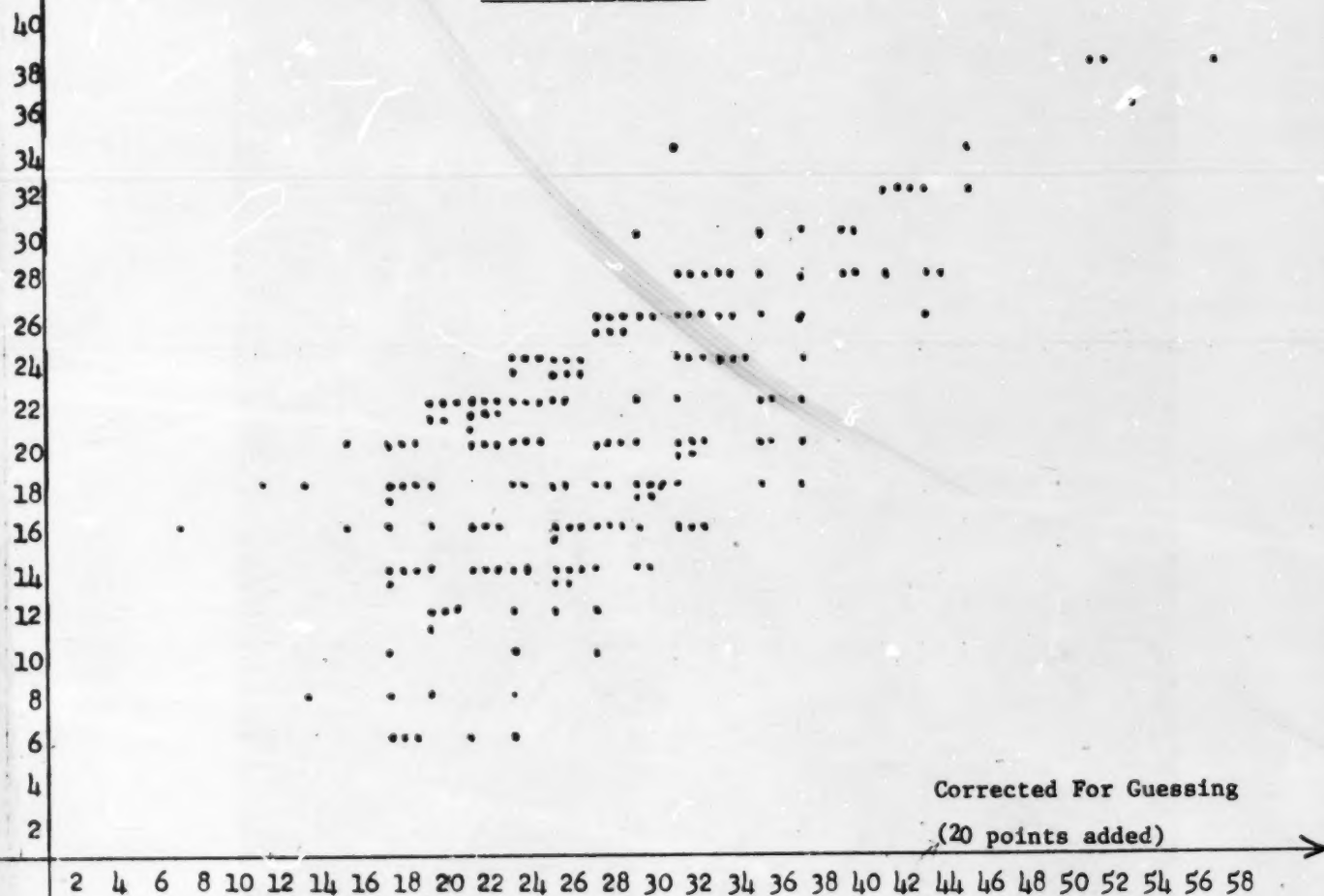
(Personal communication with Dr. Ruth Ekstrom at ETS,
April, 1975)

Appendix D

SCATTERPLOTS OF TWO METHODS OF SCORING THE SR-O TESTS

Not Corrected
For Guessing

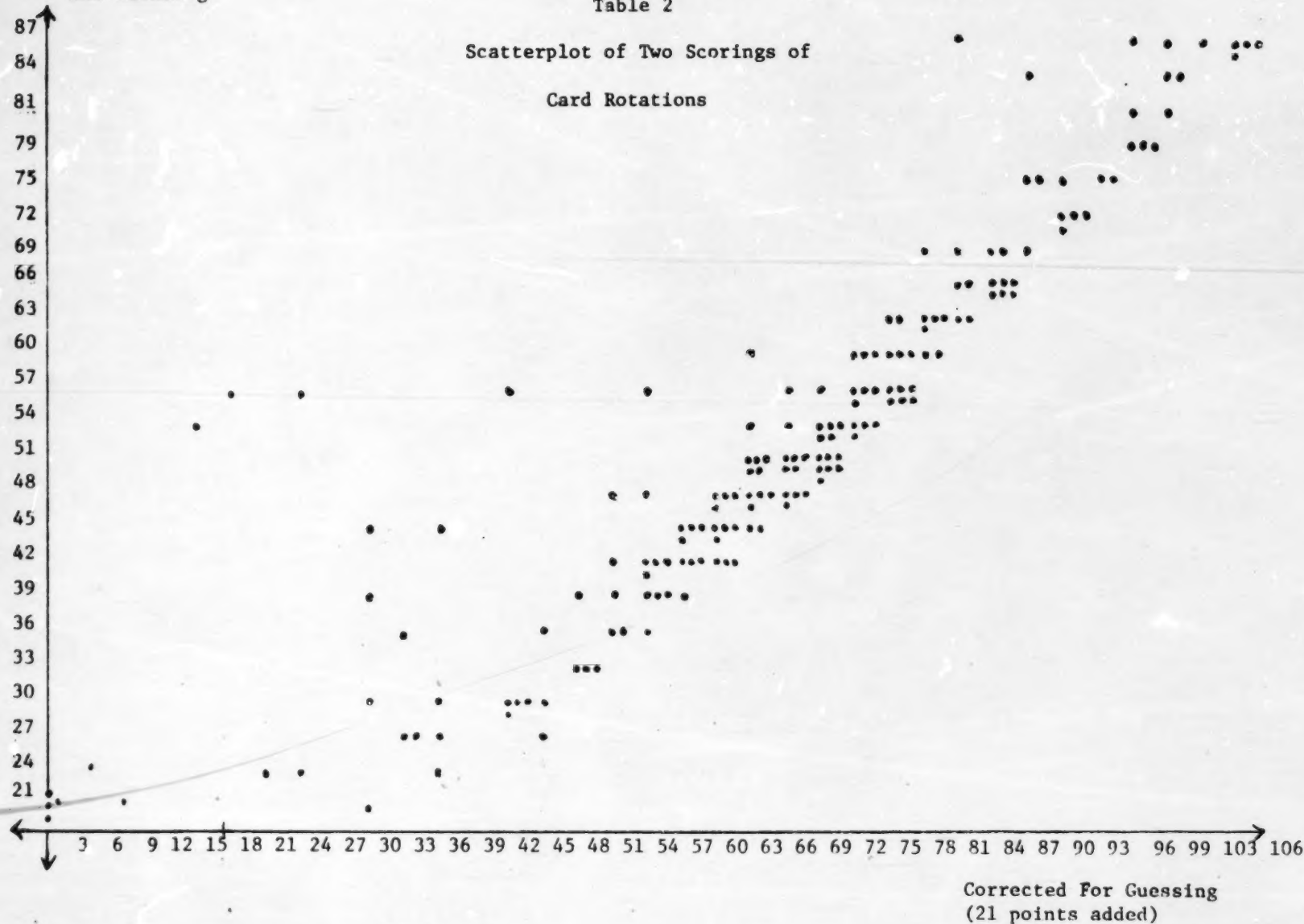
Table 1
Scatterplot of Two Scorings of
Cubes Comparison



Not Corrected
For Guessing

Table 2

Scatterplot of Two Scorings of
Card Rotations



Appendix E
CORRELATION MATRICES

C O R R E L A T I O N M A T R I X

VARIABLE NO. NO. NAME	1	2	3	4	5	6	7	8	9	10	11	12
	10	FORWARD	MIDFIGS	DATSPN	PSSOLA	PSSOLB	PSSOLC	PSSOLT	DIAGNVA	DIAGRMB	DIAGRMC	DIAGRMT
1 10	1.000											
2 FORWARD	.203	1.000										
3 MIDFIGS	.206	.045	1.000									
4 DATSPN	.307	.376	.093	1.000								
5 PSSOLA	.350	.173	.138	.277	1.000							
6 PSSOLB	.433	.181	.163	.438	.535	1.000						
7 PSSOLC	.544	.064	.090	.221	.438	.638	1.000					
8 PSSOLT	.570	.161	.153	.370	.756	.873	.861	1.000				
9 DIAGNVA	.433	-.031	.072	.219	.305	.311	.351	.389	1.000			
10 DIAGRMB	.242	.149	.057	.199	.150	.109	.191	.182	.459	1.000		
11 DIAGRMC	.112	.274	.203	.053	.013	-.074	-.011	-.030	-.114	.210	1.000	
12 DIAGRMT	.422	.075	.065	.249	.281	.258	.330	.350	.891	.804	.116	1.000
13 CREPA	.377	.052	.060	.313	.662	.445	.250	.523	.645	.213	-.133	.530
14 CREPB	.223	-.120	-.046	.097	.090	.235	.185	.209	.483	.614	-.080	.611
15 CREPC	.048	.270	.102	.097	.000	-.008	-.066	-.033	-.098	.165	.861	.092
16 CREPT	.400	.015	.013	.291	.524	.456	.276	.487	.711	.501	-.024	.721
17 PRCENTA	.253	.132	.011	.284	.655	.389	.114	.431	.305	.071	-.010	.242
18 PRCENTB	.202	-.142	-.119	.006	.062	.172	.096	.134	.268	.162	-.112	.249
19 PRCENTC	.088	.270	.102	.097	.000	-.008	-.066	-.033	-.098	.165	.861	.092
20 PRCENTT	.216	.102	-.150	.153	.442	.309	.066	.305	.257	.152	-.026	.245
21 CARDROT	.249	.321	.198	.424	.191	.350	.248	.318	.147	.223	.106	.215
22 CUCOMT	.337	.277	.104	.505	.133	.328	.323	.324	.188	-.021	.085	.128

VARIABLE NO. NO. NAME	13	14	15	16	17	18	19	20	21	22
	CREPA	CREPB	CREPC	CREPT	PRCENTA	PRCENTB	PRCENTC	PRCENTT	CARDROT	CUCOMT
13 CREPA	1.000									
14 CREPB	.245	1.000								
15 CREPC	-.099	-.026	1.000							
16 CREPT	.350	.707	.046	1.000						
17 PRCENTA	.637	.120	.045	.658	1.000					
18 PRCENTB	.172	.757	-.061	.498	.132	1.000				
19 PRCENTC	-.099	-.026	1.000	.046	.045	-.061	1.000			
20 PRCENTT	.651	.502	.352	.739	.761	.542	.752	1.000		
21 CARDROT	.173	.255	.367	.251	.169	.144	.067	.183	1.000	
22 CUCOMT	.127	.041	.032	.106	.107	.095	.032	.062	.376	1.000

NUMBER OF OBSERVATIONS 87

END OF ANALYSIS

CORRELATION MATRIX

VARIABLE NO. NO. NAME	1	2	3	4	5	6	7	8	9	10	11	12
	IC	FORMED	HIDFIGS	DATSPN	PSSOLA	PSSOLB	PSSOLC	PSSOLT	DIAGNMA	DIAGRMB	DIAGRMC	DIAGRMT
1 IC	1.000											
2 FORMED	.277	1.000										
3 HIDFIGS	.097	.159	1.000									
4 DATSPN	.357	.363	.104	1.000								
5 PSSOLA	.420	.235	.246	.384	1.000							
6 PSSOLB	.463	.336	.272	.467	.469	1.000						
7 PSSOLC	.455	.244	.266	.429	.544	.519	1.000					
8 PSSOLT	.554	.334	.319	.523	.773	.624	.859	1.000				
9 DIAGNMA	.428	.123	.150	.392	.362	.333	.255	.378	1.000			
10 DIAGRMB	.309	.137	.068	.245	.140	.359	.128	.260	.397	1.000		
11 DIAGRMC	.167	-.067	.043	.127	.086	.177	.055	.104	.230	.136	1.000	
12 DIAGRMT	.459	.135	.144	.403	.336	.399	.247	.394	.931	.685	.351	1.000
13 CREPA	.441	.227	.299	.426	.702	.466	.373	.598	.704	.241	.181	.643
14 CREPB	.471	.135	.084	.315	.188	.538	.277	.422	.385	.525	.086	.495
15 CREPC	.210	-.127	.054	.163	.195	.105	.108	.157	.236	.134	.820	.332
16 CREPT	.494	.191	.263	.431	.583	.572	.365	.603	.658	.406	.287	.713
17 PRCENTA	.366	.204	.233	.292	.684	.380	.383	.560	.314	.064	.037	.266
18 PRCENTB	.251	.066	.085	.205	.138	.384	.273	.337	.227	.130	.078	.229
19 PRCENTC	.210	-.127	.054	.163	.195	.105	.108	.157	.236	.134	.820	.332
20 PRCENTT	.389	.162	.214	.339	.529	.502	.410	.576	.281	.166	.109	.267
21 CAROMOT	.251	.416	.166	.366	.333	.400	.228	.385	.186	.044	-.166	.136
22 CUCOMT	.366	.384	.101	.509	.398	.491	.277	.468	.313	.088	.048	.275

VARIABLE NO. NO. NAME	13	14	15	16	17	18	19	20	21	22
	CREPA	CREPB	CREPC	CREPT	PRCENTA	PRCENTB	PRCENTC	PRCENTT	CAROMOT	CUCOMT
13 CREPA	1.000									
14 CREPB	.389	1.000								
15 CREPC	.294	.171	1.000							
16 CREPT	.904	.713	.427	1.000						
17 PRCENTA	.758	.295	.146	.660	1.000					
18 PRCENTB	.318	.812	.151	.580	.357	1.000				
19 PRCENTC	.294	.171	1.000	.427	.146	.151	1.000			
20 PRCENTT	.472	.673	.254	.778	.809	.787	.254	1.000		
21 CAROMOT	.254	.246	-.133	.289	.284	.230	-.133	.286	1.000	
22 CUCOMT	.420	.258	.154	.426	.278	.169	.154	.296	.332	1.000

NUMBER OF OBSERVATIONS 83

END OF ANALYSIS

75/09/12

MUN 01

PAGE 9

SUBSET 1

CORRELATION MATRIX

VARIABLE NO.	1	2	3	4	5	6	7	8	9	10	11	12
NO. NAME	1	FORMHD	HIFIGS	DATSPH	PSSOLA	PSSOLB	PSSOLC	PSSOLT	DIAGNPA	DIAGNMB	DIAGNMC	DIAGNMD
1 IC	1.000											
2 FORMHD	.221	1.000										
3 HIFIGS	.144	.079	1.000									
4 DATSPH	.331	.364	.096	1.000								
5 PSSOLA	.349	.193	.195	.328	1.000							
6 PSSOLB	.460	.235	.229	.451	.499	1.000						
7 PSSOLC	.514	.157	.161	.324	.487	.569	1.000					
8 PSSOLT	.540	.234	.244	.446	.763	.645	.858	1.000				
9 DIAGNPA	.425	-.002	.134	.300	.326	.323	.271	.367	1.000			
10 DIAGNMB	.274	.097	.076	.212	.143	.234	.140	.210	.444	1.000		
11 DIAGNMC	.146	.055	.106	.091	.054	.050	.017	.046	.135	.183	1.000	
12 DIAGNMD	.435	.044	.140	.314	.300	.334	.254	.356	.917	.750	.285	1.000
13 CREPA	.411	.101	.221	.365	.665	.456	.306	.552	.689	.244	.094	.607
14 CREPB	.316	-.049	.046	.199	.137	.399	.209	.308	.450	.585	.051	.570
15 CREPC	.162	.032	.078	.130	.111	.065	.028	.077	.137	.160	.835	.258
16 CREPT	.447	.053	.165	.314	.537	.519	.298	.531	.715	.461	.208	.727
17 PRCENTA	.312	.143	.141	.286	.667	.385	.243	.493	.315	.079	.027	.263
18 PRCENTB	.227	-.083	.002	.095	.096	.279	.167	.225	.262	.169	.015	.260
19 PRCENTC	.162	.032	.078	.130	.111	.065	.028	.077	.137	.160	.835	.258
20 PRCENTT	.304	.111	.063	.241	.482	.410	.228	.436	.281	.173	.066	.281
21 CARDROT	.250	.348	.178	.396	.258	.374	.236	.350	.166	.141	-.050	.172
22 CUCOMT	.347	.330	.095	.507	.258	.406	.305	.395	.230	.012	.049	.176

VARIABLE NO. NO. NAME	13	14	15	16	17	18	19	20	21	22
	CREPA	CREPB	CREPC	CREPT	PRCENTA	PRCENTB	PRCENTC	PRCENTT	CARDROT	CUCOMT
13 CREPA	1.000									
14 CREPB	.355	1.000								
15 CREPC	.174	.114	1.000							
16 CREPT	.748	.715	.315	1.000						
17 PRCENTA	.748	.221	.110	.655	1.000					
18 PRCENTB	.260	.784	.071	.542	.245	1.000				
19 PRCENTC	.174	.114	1.000	.315	.110	.071	1.000			
20 PRCENTT	.440	.594	.177	.752	.786	.647	.177	1.000		
21 CARDROT	.237	.247	-.047	.266	.225	.205	-.047	.233	1.000	
22 CUCOMT	.272	.129	.001	.258	.184	.116	.091	.166	.354	1.000

NUMBER OF OBSERVATIONS 170

END OF ANALYSIS

National Evaluation Committee

Francis S. Chase, Chairman
Emeritus Professor
University of Chicago
Helen Bain
Past President
National Education Association
Lyle Bourne
Professor
University of Colorado
Sue Buel
National Evaluation Committee
Roald F. Campbell
Emeritus Professor
The Ohio State University
George E. Dickson
Dean, College of Education
University of Toledo

Larry R. Goulet
Professor
University of Illinois
Chester W. Harris
Professor
University of California - Santa Barbara
William G. Katzenmeyer
Professor
Duke University
Barbara Thompson
Superintendent of Public Instruction
State of Wisconsin
Joanna Williams
Professor
Teachers College
Columbia University

University Advisory Committee

John R. Palmer, Chairman
Dean School of Education
William R. Bush
Deputy Director
R & D Center
David E. Cronon
Dean
College of Letters and Science
Diane H. Eich
Specialist
R & D Center
Evelyn L. Hoekenga
Coordinator
R & D Center
Dale D. Johnson
Associate Professor
Curriculum and Instruction
Herbert J. Klausmeier
Member of the Associated Faculty
R & D Center

James M. Lipham
Member of the Associated Faculty
R & D Center
Wayne R. Otto
Associate Director
R & D Center
Richard A. Rossmiller
Director
R & D Center
Elizabeth J. Simpson
Dean
School of Family Resources
and Consumer Sciences
Len Van Ess
Associate Vice Chancellor
University of Wisconsin - Madison

Associated Faculty

Vernon L. Allen
Professor
Psychology
B. Dean Bowles
Professor
Educational Administration
Thomas P. Carpenter
Assistant Professor
Curriculum and Instruction
Marvin J. Fruth
Professor
Educational Administration
John G. Harvey
Professor
Mathematics
Curriculum and Instruction
Frank H. Hooper
Professor
Child Development
Herbert J. Klausmeier
V A C. Henmon Professor
Educational Psychology
Joseph T. Lawton
Assistant Professor
Child Development

Joel R. Levin
Professor
Educational Psychology
L. Joseph Lins
Professor
Institutional Studies
James M. Lipham
Professor
Educational Administration
Donald N. McIsaac
Professor
Educational Administration
Gerald Nadler
Professor
Industrial Engineering
Wayne R. Otto
Professor
Curriculum and Instruction
Robert G. Petzold
Professor
Music
Curriculum and Instruction

Thomas S. Popkewitz
Assistant Professor
Curriculum and Instruction
Thomas A. Romberg
Professor
Curriculum and Instruction
Richard A. Rossmiller
Professor
Educational Administration
Dennis W. Spuck
Assistant Professor
Educational Administration
Michael J. Subkoviak
Assistant Professor
Educational Psychology
Richard L. Venezky
Professor
Computer Sciences
J. Fred Weaver
Professor
Curriculum and Instruction
Larry M. Wilder
Assistant Professor
Child Development